OF COMMERCE
andards

**NBS Technical Note 1246**

*A New Statistical Model for the Calibration of Force Sensors*

*Charles P. Reeve*

NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NE
NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NE
NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NE
NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NE
NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NE
NBS NBS NBS *National Bureau of Standards* NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NE
NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NB
VBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
S NBS NBS NBS NBS NBS NBS NBS NBS NBS NB

*T*he National Bureau of Standards[1] was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research to assure international competitiveness and leadership of U.S. industry, science and technology. NBS work involves development and transfer of measurements, standards and related science and technology, in support of continually improving U.S. productivity, product quality and reliability, innovation and underlying science and engineering. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, the Institute for Computer Sciences and Technology, and the Institute for Materials Science and Engineering.

## *The National Measurement Laboratory*

Provides the national system of physical and chemical measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; provides advisory and research services to other Government agencies; conducts physical and chemical research; develops, produces, and distributes Standard Reference Materials; provides calibration services; and manages the National Standard Reference Data System. The Laboratory consists of the following centers:

- Basic Standards[2]
- Radiation Research
- Chemical Physics
- Analytical Chemistry

## *The National Engineering Laboratory*

Provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

- Applied Mathematics
- Electronics and Electrical Engineering[2]
- Manufacturing Engineering
- Building Technology
- Fire Research
- Chemical Engineering[3]

## *The Institute for Computer Sciences and Technology*

Conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following divisions:

- Information Systems Engineering
- Systems and Software Technology
- Computer Security
- Systems and Network Architecture
- Advanced Computer Systems

## *The Institute for Materials Science and Engineering*

Conducts research and provides measurements, data, standards, reference materials, quantitative understanding and other technical information fundamental to the processing, structure, properties and performance of materials; addresses the scientific basis for new advanced materials technologies; plans research around cross-cutting scientific themes such as nondestructive evaluation and phase diagram development; oversees Bureau-wide technical programs in nuclear reactor radiation research and nondestructive evaluation; and broadly disseminates generic technical information resulting from its programs. The Institute consists of the following Divisions:

- Ceramics
- Fracture and Deformation[3]
- Polymers
- Metallurgy
- Reactor Radiation

# NBS Technical Note 1246

# A New Statistical Model for the Calibration of Force Sensors

Charles P. Reeve

Center for Applied Mathematics
National Engineering Laboratory
National Bureau of Standards
Gaithersburg, MD 20899

June 1988

# Contents

# List of Figures

# List of Tables

# A New Statistical Model for the Calibration of Force Sensors

Charles P. Reeve

### Abstract

The National Bureau of Standards has been calibrating force sensors for many years. The objective in these calibrations is to determine the functional relationship between the applied load and the sensor response. In a typical calibration several runs are made in which identical sequences of known loads are applied to the sensor. The sensor is rotated in the loading machine between runs. The previous method of analysis incorporated a quadratic polynomial model which was fit to the pooled data. The new method presented here fits separate polynomials to the mean data and between-run differences. The 'best' degrees of these polynomials are automatically determined by algorithms which incorporate statistical tests. As a result, error contributions from several sources are quantized. Methods for computing confidence intervals for the 'true' sensor response and for a new observation are given, and methods of inverse prediction (measurement of an unknown load) based on these intervals are illustrated.

*Key words*: calibration curve; confidence interval; force sensor; inverse prediction; least squares polynomial; load cell; proving ring; statistical model

## 1 Introduction

The National Bureau of Standards (NBS) has been calibrating force measuring devices for many years. The current status of these calibrations has been described by Mitchell [13]. In that paper the generic term 'force sensor' is used to include '... instruments and systems that are variously referred to by names such as load cell, proving ring, force gage, force link, force transducer, load ring, ring dynamometer, compression dynamometer, tension dynamometer, and crane scale.' Hereafter the generic term will be used except in reference to a specific type of sensor. These sensors are normally used to measure forces in the 10 to 1,000,000 pound-force (lbf) range. Their response curves typically look linear to the eye as shown in figure 1.

The objective in calibrating a force sensor is to determine the functional relationship between the applied load and the sensor response. This is accomplished in the laboratory by applying a series of known loads to the sensor and observing its response on a readout

instrument. The calibration is usually complicated by the interaction of the sensor with the loading machine as discussed by Mitchell and Pontius [15]. Also, the sensor may creep or drift under sustained load (see Mitchell and Baker [14]) and it may be sensitive to atmospheric conditions such as temperature and pressure. A general discussion of the above problems and others is given by Pontius and Mitchell [19].

A series of interlaboratory comparisons of several force sensors over a seven year period has been reported by Peterson, Jenkins, and Mitchell [18] (following a preliminary report by Peterson and Bloss [17]). In general the force sensors exhibited reasonable short term stability, but there was some evidence of long term drift which indicates the need for their periodic recalibration.

The purpose of this paper is to present a new statistical model for the force sensor calibration process which takes into account errors from several known sources. The model, based on variable-degree polynomials, is designed to extract as much information as possible from data taken in the traditional manner. By quantifying the various kinds of errors, the underlying processes may be better understood as more calibration results are analyzed by the new model.

In developing the new model many sets of force sensor data were analyzed. As more data are collected and analyzed the model may need to be modified. For this reason the methods in this paper should not be construed as the ultimate answer to the force sensor calibration problem.

The mathematical and statistical aspects of the model are described in considerable detail in section 4. A nontechnical outline of the new model is given in section 3. Some preliminary discussions of other aspects of force sensor calibration are condensed from the previously referenced papers and given in section 2. The reader would do well to be familiar with those references, especially [13], beforehand.

An example with real data is given in section 5, and the results of a simulation study based on this example are reported in section 6. A question of optimality in the choice of applied loads is considered in section 7.

The FORTRAN computer program FCAL88 has been written to perform the analyses described in this paper. The program is self contained and has been successfully run on both mainframe and personal computers. A copy of the program can be made available to the user upon request.

# 2   Preliminaries

## 2.1   Previous Methods

In 1946 Wilson, Tate, and Borkowski [25] published a paper on the calibration of proving rings in which 'calibration factors', the ratio of the load to the corresponding deflection, were computed. Corresponding to these factors were 'specification limits' which served as uncertainty bounds. The factors were not quite constant with load, thus indicating a slight

nonlinearity in the proving ring response. Errors from various sources and loading sequences were also discussed.

In 1964 Hockersmith and Ku [11] developed a statistical model for proving rings in which the response (the pooled runs) was taken to be a quadratic function of the applied load. A method for computing a confidence band for the 'true' underlying response function was given. This model proved satisfactory and was incorporated into the calibration process for both proving rings and load cells.

## 2.2 Loading Sequences

Many force sensors can be calibrated in both tension and compression modes. The response is expected to be somewhat different in each mode. Due to hysteresis effects the response may also depend on whether the loads are applied in ascending or descending order. The device may thus have several distinct calibration curves.

The ASTM Standard [1], currently most frequently used at NBS, requires at least thirty measurements over the working range of the device during calibration. At NBS the calibration often consists of two runs at fifteen loads or three runs at ten loads. In the latter case the applied loads are typically at 10%, 20%, 30%, ..., 90%, and 100% of capacity. Between runs the sensor is rotated in the loading machine. This rotation allows errors from misalignment to be sampled.

It is necessary to make measurements at zero load in order to define the baseline for the readout of the sensor. Since conditions under zero load can be different than under nonzero loads, those readings have to be treated differently. The manner in which they are treated depends somewhat on the characteristics of the loading machine. Some deadweight machines require weights to be added sequentially while others require a return to zero between loads. In general a run will consist of a series of readings at zero and nonzero loads. From these readings a set of deflections can be computed for each of the nonzero loads adjusted for instrument drift to a zero baseline. The manner of adjustment depends on the theorized drift behavior of the instrument between zero loads. If the instrument is assumed not to drift then an initial zero load reading is subtracted from the following nonzero load readings. If a linear drift is assumed then nonzero load readings are adjusted proportionally according to 'before' and 'after' zero load readings. It will be assumed that all loads are equally spaced in time so that drift corrections can be easily computed. It is imperative that *the same loading sequence be used for each run.*

In [19] a quadratic curve is said to be fit to the $y - y_0$ data. This means that readings $y$ were taken at various loads and the reading of the instrument under zero load $y_0$ was subtracted from it. In cases where a particular $y_0$ value or a combination of $y_0$ values is subtracted from more than one $y$ value the resulting $y - y_0$ values are correlated. Previous fitting procedures have ignored this correlation by assuming that all the $y - y_0$ were independent. An attempt to incorporate these correlations into the new model was made. However, the computational complexities introduced were felt to outweigh the aesthetic value in 'correctly' modeling the

3

observations. As a result the $y - y_0$ values will continue to be assumed independent in the new method of analysis.

## 2.3  Sources of Error

The method of analysis in [1,11,19], discussed earlier, is to pool all the runs and fit a quadratic curve to the combined data. Upon examining hundreds of these fits to load cell data the two most often seen phenomena are a serpentine ($S$-shaped) curve in the residuals, shown in figure 3, and a fanning out of the residuals from each run as shown to some extent in figure 8. The serpentine curve is thought to be the result of *model error*, that is, the quadratic curve insufficiently modeling the actual underlying sensor characteristic. At this point there is no solid theoretical basis for stating what the response should be.

In this study many mathematical models were tried in hopes of finding a general curve which would adequately fit the vast majority of calibration data. These models include quadratic and higher polynomials, half power polynomials, negative power polynomials, and exponential, sinusoidal, logarithmic, and Bessel functions. In specific cases each of these has provided a good fit, but none has been found satisfactory in general. The most success was had with quadratic, cubic, and quartic polynomials. The cubic term was found to be the most effective way of removing the serpentine curve in many cases. The decision was thus made to base the new method on seeking higher order polynomial terms which will significantly decrease the residual error.

The fanning out of residuals is due to *between-run error*, referred to in [15] as consisting of 'zero shift' and 'load proportional' components. This error is believed to be due to changes in mechanical alignment (or misalignment) when the sensor is rotated, with the magnitude of the error being dependent on the sensor/machine combination. Ideally, the net applied load is parallel and symmetric with respect to the axis of the sensor, and the sensor response function is constant over time. In reality, however, small components of non-axisymmetric forces exist whenever a load is applied. The sensor response depends on the magnitude of those forces, the sensitivity of the sensor to those forces, and the orientation of the sensor in the loading machine. Between-run errors appear to take the form of separate polynomials of the same degree but possibly different coefficients.

A study by Mitchell, Seifarth, and Reeve [16] showed that eccentric (off-center) loading results in sinusoidal differences in response as a function of angle of orientation of the load eccentricity. By making runs with the sensor equiangularly spaced around the circle, the sinusoidal components of error tend to cancel out when the runs are averaged.

Other sources of error are generally of smaller magnitude and easier to model. The random fluctuations in the output under repeated loads of the same magnitude are another kind of error. These may be due to many separate physical processes which are ordinarily not measured during a calibration, for example, fluctuations in atmospheric conditions, vibrations, and oscillations of the dead weights. The effect of all these short term variations combine to give what looks like random 'noise' and can be called *within-run error*.

4

MEAN OATA WITH FITTEO STRAIGHT LINE
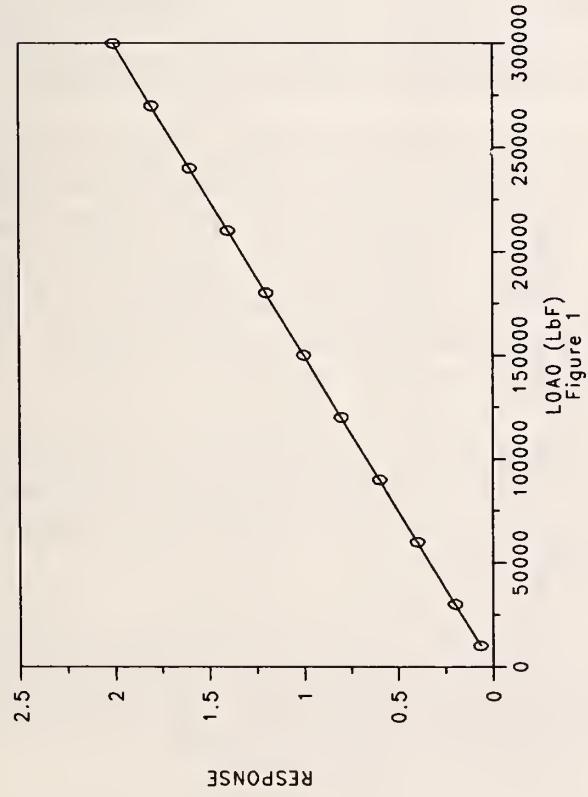
RESIDUALS FROM LINEAR FIT TO MEAN OATA
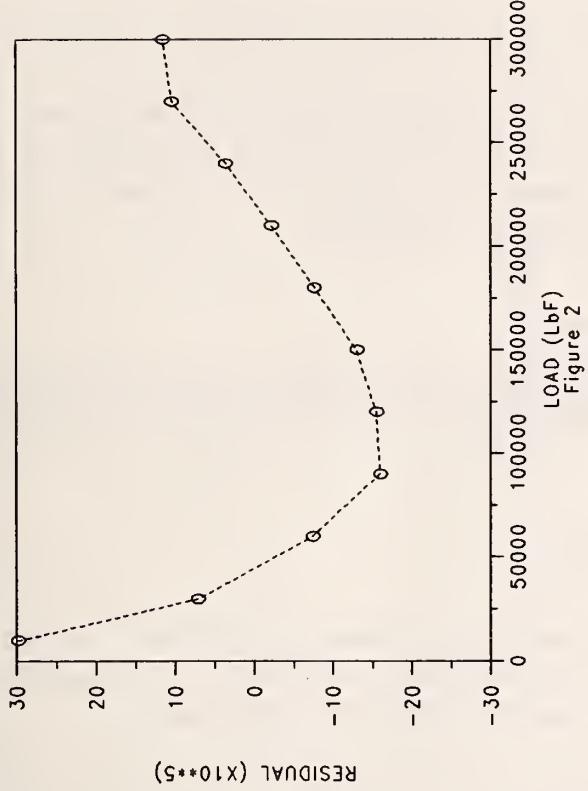
RESIOUALS FROM QUADRATIC FIT TO MEAN OATA
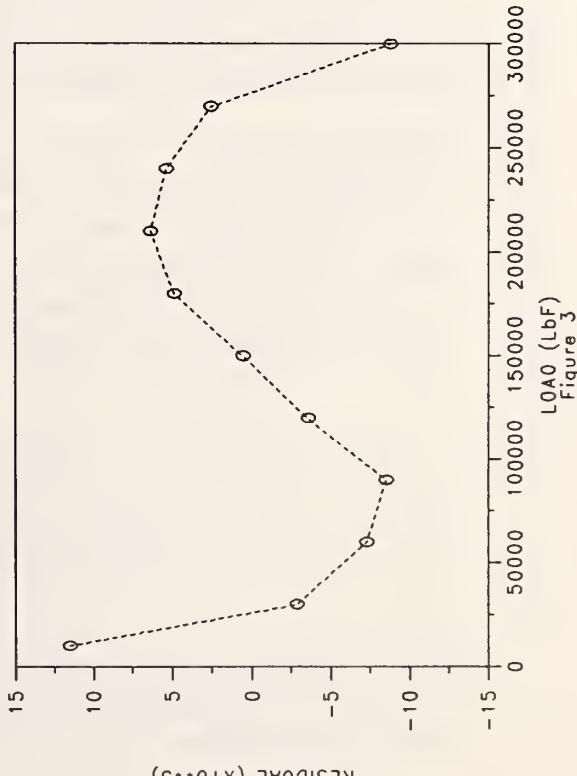
RESIOUALS FROM CUBIC FIT TO MEAN OATA

LOAD (LbF)
Figure 1

LOAD (LbF)
Figure 2

LOAO (LbF)
Figure 3

LOAO (LbF)
Figure 4

5
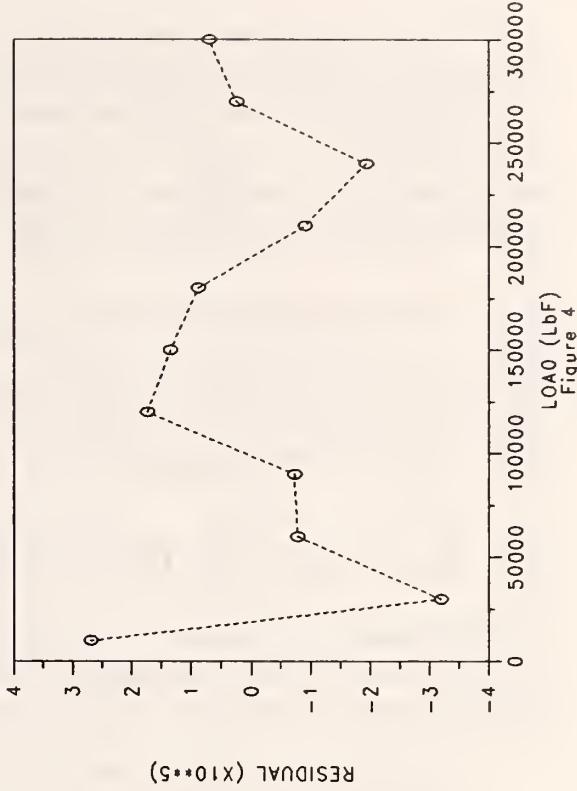
When the instrument readout is too coarse a *roundoff error* results. This type of error can easily mask the effects of other errors.

Finally, errors in the accepted dead weight values whether due to calibration error or air buoyancy effects are a source of *systematic error*. A term to account for this may be added at the end of the calibration. The value for this error at NBS is estimated not to exceed 0.002%. This term is not added in the ASTM method of analysis.

## 2.4 Preparation for Measurement

The force sensor is mounted in the loading machine in either the tension or compression mode. Several 'warm-up' or 'pre-loading' runs are then made in order to stabilize the response of the device and reduce hysteresis effects.

# 3 Features of This Method

The purpose of this section is to describe in nontechnical terms how the new method of analysis works. The reader who is interested in the details of the statistical model and analysis will find those in section 4. The five main features of the method are:

1. *A polynomial response model is fit to the mean data* (the average of all the runs). Subject to certain limitations *the degree of this polynomial is determined from the data.* The upper limit on the degree of the polynomial is set by the user. Figures 2, 3, and 4 show residuals from linear, quadratic, and cubic fits respectively to the mean data from the example in appendix B (note the different vertical scales). The cubic fit was determined to be 'best'. In this case the upper limit on the degree was five.

2. *A polynomial response model is fit to each set of difference data* (the difference between the individual runs and the mean data). Subject to certain limitations *the degree of these polynomials is automatically determined from the data.* The degree of each polynomial is constrained to be the same, and as before the upper limit on that degree is set by the user. For the example the difference data is shown in figures 5, 6, and 7 with a straight line fit to each run. The combined data are shown in figure 8. The linear fit was determined to be 'best'. The upper limit on the degree was three.

3. *A rough error budget is computed for each of the five sources of error* discussed in section 2.3. This summary may prove to be useful in correlating the error structure of the data with the underlying physical processes. The error budget for the example is shown in table 2 in section 5 in the $k^* = 3$ column.

4. *Three types of confidence intervals are computed* and their proper usage is discussed. These intervals provide the user with a realistic estimate of the uncertainty associated with the calibrated force sensor response function, and they allow the realistic uncertainty of further use, such as inverse prediction, to be computed.
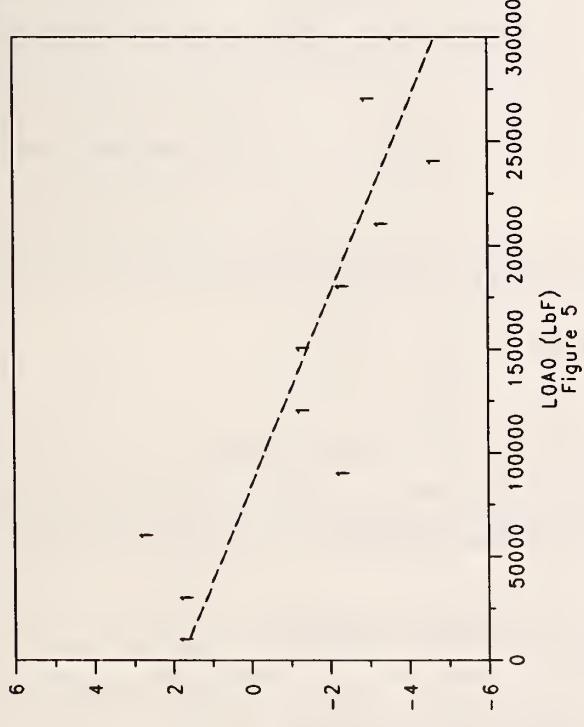
6

DIFFERENCE OATA WITH FITTEO POLYNOMIAL - RUN 2

LOAO (LbF)
Figure 6

RESPONSE (X10**5)

COMBINEO OIFFERENCE OATA WITH FITTEO POLYNOMIALS

LOAO (LbF)
Figure 8

RESPONSE (X10**5)

OIFFERENCE OATA WITH FITTEO POLYNOMIAL - RUN 1

LOAO (LbF)
Figure 5

RESPONSE (X10**5)

OIFFERENCE OATA WITH FITTEO POLYNOMIAL - RUN 3

LOAO (LbF)
Figure 7

RESPONSE (X10**5)

7
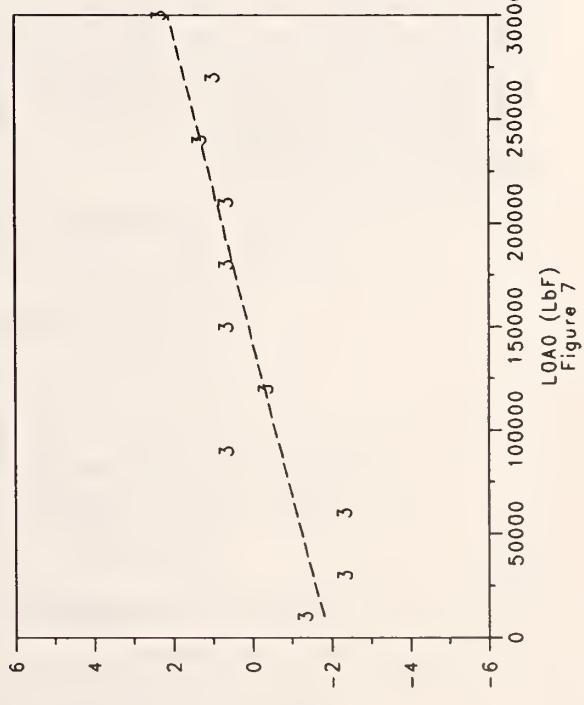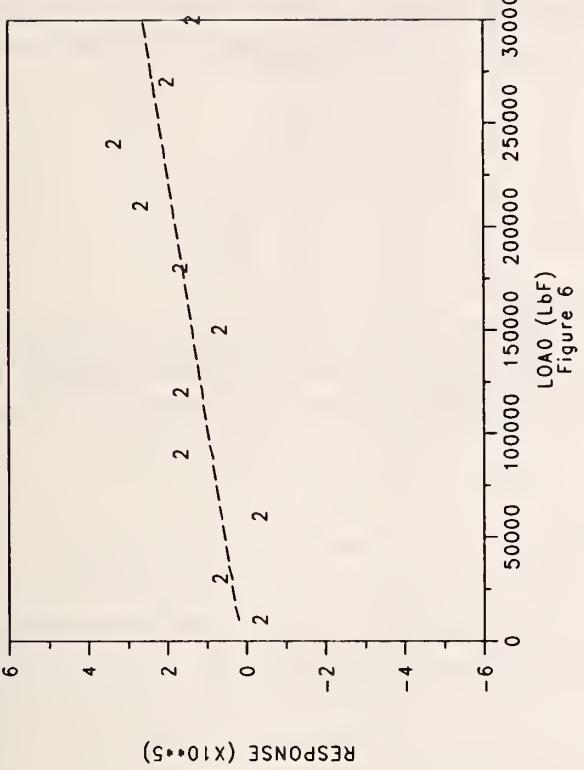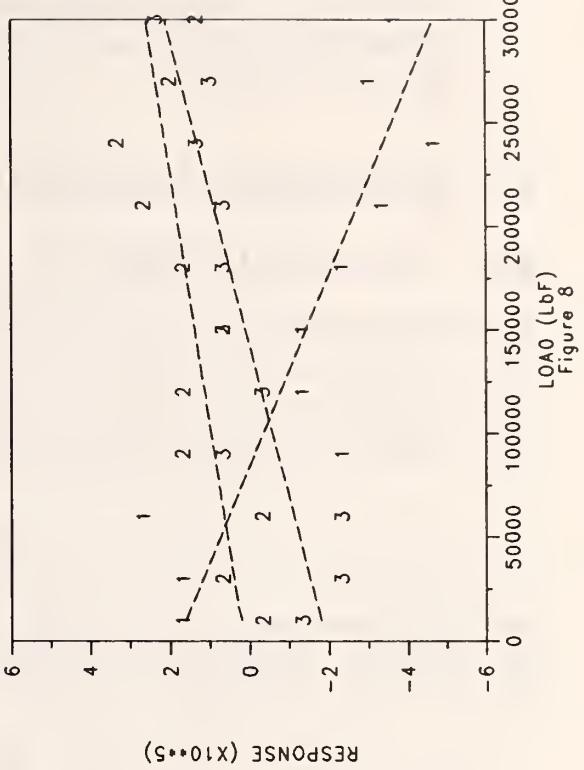
5. *Three types of inverse prediction are discussed* in which unknown loads are measured by the calibrated force sensor. The uses of these methods are graphically illustrated in figures 9, 10, and 11.

# 4 Statistical Analysis

## 4.1 Statistical Model

The sensor response to axial forces may be considered fixed throughout the calibration, but the response to non-axial forces must be considered random because it changes unpredictably whenever the sensor is re-oriented in the loading machine. The fixed response to axial loads $L_i$ is modeled by the $k$-degree polynomial

$$p_k(L_i) = \sum_{t=0}^{k} \beta_t L_i^t \qquad (1)$$

where $k$ is usually two or three. During the $j^{th}$ orientation of the sensor ($j^{th}$ run), the random response to non-axial forces is modeled by the $h$-degree polynomial

$$q_{hj}(L_i) = \sum_{t=0}^{h} u_{tj} L_i^t$$

where $h$ is usually zero or one and $j$ indexes the runs. Though assumed to be drawn randomly from a population, the $\{u_{tj} | t = 0, 1, \ldots, h\}$ are treated as fixed effects for the duration of the $j^{th}$ run. In some cases the random effects may prove to be not significant, in which case $h$ is taken to be $-1$ and the $\{u_{tj}\}$ do not exist.

The following notation will be used in the statistical analysis which incorporates these polynomial models. Let

$$
\begin{aligned}
n \;&=\; \text{the number of distinct nonzero loads applied per run,} \\
r \;&=\; \text{the number of runs (sequences of } n \text{ loads),} \\
L_i \;&=\; \text{the } i^{th} \text{ nonzero load in the loading sequence} \\
&\qquad (i = 1, 2, \ldots, n), \\
y_{ij} \;&=\; \text{the observed (net) value of the instrument output} \\
&\qquad \text{when load } L_i \text{ is applied during run } j \ (j = 1, 2, \ldots, r), \\
\beta_t \;&=\; \text{the coefficient of } L_i^t \text{ in the fixed response} \\
&\qquad (t = 0, 1, \ldots, k), \\
u_{tj} \;&=\; \text{the coefficient of } L_i^t \text{ in the random response} \\
&\qquad (t = 0, 1, \ldots, h; j = 1, 2, \ldots, r), \\
k \;&=\; \text{the degree of the fixed response polynomial,}
\end{aligned}
$$

8

$$
\begin{aligned}
h &= \text{the degree of the random response polynomials,} \\
k_{min} &= \text{the lower bound on } k \text{ (usually 1),} \\
h_{min} &= \text{the lower bound on } h \text{ (usually } -1\text{),} \\
k_{max} &= \text{the upper bound on } k, \\
h_{max} &= \text{the upper bound on } h, \text{ and} \\
\epsilon_{ij} &= \text{the random 'noise' in the instrument output} \\
&\quad \text{when the } i^{th} \text{ load is applied during the } j^{th} \text{ run.}
\end{aligned}
$$

The statistical model then takes the form

$$
\begin{aligned}
y_{ij} &= p_k(L_i) + q_{hj}(L_i) + \epsilon_{ij} \\
&= \beta_0 + \beta_1 L_i + \ldots + \beta_k L_i^k + u_{0j} + u_{1j} L_i + \ldots + u_{hj} L_i^h + \epsilon_{ij} \qquad (2)
\end{aligned}
$$

where the polynomial degrees $k$ and $h$ $(k \geq h)$ are to be determined from the data by statistical tests. The $\{\epsilon_{ij}\}$ are assumed to be independent error values from a distribution with mean zero and variance $\sigma_e^2$. (Recall from section 2.2 that the $y_{ij}$ are assumed to be independent even though correlations *may* have been introduced through corrections for instrument drift). The vectors of random coefficients $(u_{0j}\ u_{1j}\ \ldots\ u_{hj})^T$ $(j = 1, 2, \ldots, r)$ are assumed to be independent with expectation $(0\ 0\ \ldots\ 0)^T$ and variance $\Sigma_u$, an $(h+1) \times (h+1)$ symmetric positive definite matrix. (The '$T$' superscript indicates matrix transposition.)

The 'best' degree polynomial fit to data may be determined either by starting with a low degree polynomial and adding higher order terms which are significant, or by starting with a high degree polynomial and deleting higher order terms which are not significant. After due consideration the method of 'deleting terms', as recommended by Hager and Antle [9], was determined to be the most appropriate for computing the 'best' values of both $k$ and $h$ (also see Hoel [12]). This requires the user to establish the upper bounds $k_{max}$ and $h_{max}$ (on $k$ and $h$ respectively) for the initial fits.

A convenient way to analyze model (2) is to work with the independent sets of observations $\{\bar{y}_i\}$ and $\{z_{ij}\}$ derived from the $\{y_{ij}\}$ by

$$
\begin{aligned}
\bar{y}_i &= \sum_{j=1}^{r} y_{ij}/r \quad (i = 1, 2, \ldots, n) \text{ and} \\
z_{ij} &= y_{ij} - \bar{y}_i \quad (j = 1, 2, \ldots, r).
\end{aligned}
$$

These sets can be identified as the *mean* data and the *difference* data respectively.

In matrix notation *the model for the mean data* is

$$
\bar{\mathbf{y}} = \mathbf{X}\beta + \mathbf{X}_1\bar{\mathbf{u}} + \bar{\epsilon} \qquad (3)
$$

where

$$
\bar{\mathbf{y}} = (\bar{y}_1\ \bar{y}_2\ \cdots\ \bar{y}_n)^T,
$$

9

$$\beta = (\beta_0 \ \beta_1 \ \ldots \ \beta_k)^T,$$

$$\bar{u} = (\bar{u}_0 \ \bar{u}_1 \ \ldots \ \bar{u}_h)^T \ \text{with} \ \bar{u}_t = \sum_{j=1}^{r} u_{tj}/r,$$

$$\bar{\epsilon} = (\bar{\epsilon}_1 \ \bar{\epsilon}_2 \ \ldots \ \bar{\epsilon}_n)^T \ \text{with} \ \bar{\epsilon}_t = \sum_{j=1}^{r} \epsilon_{tj}/r,$$

and

$$\mathbf{X} = (\mathbf{X}_1 \ \mathbf{X}_2) = \begin{pmatrix} 1 & L_1 & \ldots & L_1^h & | & L_1^{h+1} & \ldots & L_1^k \\ 1 & L_2 & \ldots & L_2^h & | & L_2^{h+1} & \ldots & L_2^k \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ 1 & L_n & \ldots & L_n^h & | & L_n^{h+1} & \ldots & L_n^k \end{pmatrix}$$

In statistical terminology (3) is a 'mixed' model because of the presence of fixed ($\beta$) and random ($u, \epsilon$) effects. It follows that $E(\bar{y}) = \mathbf{X}\beta$ and $Var(\bar{y}) = (\mathbf{X}_1 \Sigma_u \mathbf{X}_1^T + \sigma_e^2 \mathbf{I})/r = \mathbf{V}$ where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{V}$ is an $n \times n$ covariance matrix which is positive definite when $\sigma_e^2 > 0$ [$E(\cdot)$ and $Var(\cdot)$ denote the statistical expectation and variance]. Giesbrecht and Burns [8] show that, when estimates of $\Sigma_u$ and $\sigma_e^2$ are available, the generalized least squares estimate of the fixed effects is $\hat{\beta} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}}^{-1} \bar{y}$ with $\widehat{Var(\hat{\beta})} = \hat{\Sigma}_{\hat{\beta}} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1}$ where $\hat{\mathbf{V}} = (\mathbf{X}_1 \hat{\Sigma}_u \mathbf{X}_1^T + \hat{\sigma}_e^2 \mathbf{I})/r$. In the 'large sample' case the estimate $\hat{\beta}$ is known to have a smaller variance than the ordinary least squares estimate. (Methods for testing hypotheses, computing confidence intervals, and approximating degrees of freedom are also given in [8]. For more discussion of the mixed model see Cunningham and Henderson [5], Henderson [10], and Searle [21;22, ch. 9-10].)

In section 4.2 it will be shown that the variance estimate $\hat{\Sigma}_u$ with $r-1$ degrees of freedom is obtainable from the difference data. Current practice at NBS is to make three runs, thus $\hat{\Sigma}_u$ is based on only two degrees of freedom. Although the number of runs could be increased some, it is not feasible to make $r-1$ 'large'. Since bad estimates of $\beta$ could result from using $\hat{\Sigma}_u$ based on very few degrees of freedom, the above generalized least squares approach will be set aside in favor of ordinary (unweighted) least squares.

For the purpose of computation the $\{u_{tj}\}$ are treated as fixed effects, thus *model (3) for the mean data* can be re-expressed as

$$\bar{y} = \mathbf{X}\theta + \bar{\epsilon} \tag{4}$$

where

$$\theta = (\beta_0 + \bar{u}_0 \ \ \beta_1 + \bar{u}_1 \ \ldots \ \beta_h + \bar{u}_h \ \ \beta_{h+1} \ \ \beta_{h+2} \ \ldots \ \beta_k)^T$$
$$= (\theta_0 \ \theta_1 \ \ldots \ \theta_k)^T \tag{5}$$

and $\bar{\epsilon}$ is as before. In this formulation the mean effect $\bar{u}_t$, treated as a fixed effect, gets 'absorbed' into the effect $\beta_t$ ($t = 1, 2, \ldots, h$) thus only the net effect $\theta_t$ can be estimated. It

follows that

$$\begin{aligned} \mathrm{E}(\bar{\mathbf{y}}) &= \mathbf{X}\theta, \text{ and} \\ \mathrm{Var}(\bar{\mathbf{y}}) &= \mathbf{I}\sigma_e^2/r. \end{aligned} \tag{6}$$

This model will be analyzed in section 4.3.

The *model for the difference data* is

$$\begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_r \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_1 \end{pmatrix} \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_r \end{pmatrix} + \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_r \end{pmatrix} \tag{7}$$

where

$$\mathbf{z}_j = \begin{pmatrix} z_{1j} \\ z_{2j} \\ \vdots \\ z_{nj} \end{pmatrix}, \quad \mathbf{d}_j = \begin{pmatrix} u_{0j} - \bar{u}_0 \\ u_{1j} - \bar{u}_1 \\ \vdots \\ u_{hj} - \bar{u}_h \end{pmatrix}, \quad \mathbf{e}_j = \begin{pmatrix} \epsilon_{1j} - \bar{\epsilon}_1 \\ \epsilon_{2j} - \bar{\epsilon}_2 \\ \vdots \\ \epsilon_{nj} - \bar{\epsilon}_n \end{pmatrix},$$

and the zero matrices ($\mathbf{0}$) are $n \times (h+1)$. Treating the differences $\mathbf{d}_j$ as fixed effects, it follows that

$$\begin{aligned} \mathrm{E}(\mathbf{z}_j) &= \mathbf{X}_1 \mathbf{d}_j \text{ and} \\ \mathrm{Var}(\mathbf{z}_j) &= \mathbf{I}\sigma_e^2(r-1)/r \quad (j = 1, 2, \ldots, r). \end{aligned} \tag{8}$$

This model is analyzed in the next section.

## 4.2   Analysis of Difference Data

Applying the usual method of unweighted least squares (see Cameron [2, pp. 24-26]) to (8), for a given $h \geq 0$, yields the least squares estimates

$$\hat{\mathbf{d}}_j = (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{z}_j \quad (j = 1, 2, \ldots, r). \tag{9}$$

Recall from section 4.1 that the between-run effects modeled by $\{u_{tj}\}$ may not always be significant. When that is the case there are no $\{\mathbf{d}_j\}$ to estimate, since model (7) is degenerate, and $h$ is taken to be $-1$.

The total residual sum of squares, incorporating the $r$ systems in (7) and (8), is

$$\mathrm{SS}_e(h) = \begin{cases} \sum_{j=1}^r (\mathbf{z}_j - \mathbf{X}_1 \hat{\mathbf{d}}_j)^T (\mathbf{z}_j - \mathbf{X}_1 \hat{\mathbf{d}}_j) & (h \geq 0) \\ \sum_{j=1}^r \mathbf{z}_j^T \mathbf{z}_j & (h = -1) \end{cases} \tag{10}$$

so that

$$\mathrm{E}[\mathrm{SS}_e(h)] = (n - h - 1)(r - 1)\sigma_e^2. \tag{11}$$

11

Note that the factor $r-1$ appears in (11) instead of $r$ since the means $\{\bar{\epsilon}_i\}$ were subtracted from the error values $\{\epsilon_{ij}\}$ in (7).

The following algorithm is used to determine from the difference data the 'best' degree, $h^*$, of the polynomials $q_{hj}(L_i)$ modeling the $\{u_{tj}\}$ effects. Statistical hypothesis testing at a level of significance $\alpha_h$ is used to determine whether, for a given $h$, the terms $\{u_{hj}L_i^h | j = 1, 2, \ldots, r\}$ are needed. The $h_{max}$-degree model is fit first, then insignificant higher degree terms are dropped $r$ at a time. The constraint $h_{min} \leq h^* \leq h_{max}$ is applied where $h_{min} = -1$.

### Algorithm H.

1. Set $h = h_{max}$.

2. If $h = h_{min}$ then step 5. Else,

3. Set $\nu_e = (n - h - 1)(r - 1)$ and compute the statistic
$$F_h = \frac{[\text{SS}_e(h-1) - \text{SS}_e(h)]/(r-1)}{\text{SS}_e(h)/\nu_e}$$
which has the $F(r-1, \nu_e)$ distribution under the hypothesis $u_{h1} = u_{h2} = \ldots = u_{hr} = 0$. Use (9) and (10) as appropriate.

4. If $G_{F(r-1,\nu_e)}(F_h) < 1 - \alpha_h$ set $h = h - 1$ then step 2 (the left hand side of the inequality is the cumulative distribution function of the $F$ distribution with parameters $r-1$ and $\nu_e$ as defined in appendix $A$). Else,

5. Set $h^* = h$,
$$\nu_e = (n - h^* - 1)(r - 1) \tag{12}$$
$$\hat{\sigma}_e^2 = \text{SS}_e(h^*)/\nu_e, \tag{13}$$
and, if $h^* \geq 0$,
$$\hat{\Sigma}_u = \sum_{j=1}^{r} \hat{\mathbf{d}}_j \hat{\mathbf{d}}_j^T/(r-1). \tag{14}$$

6. Done.

When $h^* \geq 0$ the estimated mean squared error for between-run effects is computed from (10) and (11) to be
$$\hat{\sigma}_b^2 = \frac{\text{SS}_e(-1) - \text{SS}_e(h^*)}{(r-1)(h^*+1)} \tag{15}$$
The statistic $F_b = \hat{\sigma}_b^2/\hat{\sigma}_e^2$ then has the $F[(r-1)(h^*+1), \nu_e]$ distribution under the hypothesis $\{u_{tj} = 0 | t = 0, 1, \ldots, h^*; j = 1, 2, \ldots, r\}$. It may seem strange to calculate this statistic when the fact that $h \geq 0$ already implies that the between-run effects are significant as determined by the statistic $F_h$ in step 3 of algorithm $H$. The usefulness of $F_b$ is that it tests the significance of the $r(h^* + 1)$ nonzero $\{u_{tj}\}$ effects as a whole whereas $F_h$ tests only $r$ of the $\{u_{tj}\}$ at a time. The level of significance of $F_b$ may then be displayed as the significance of the between-run effects.

## 4.3 Analysis of Mean Data

Applying the method of unweighted least squares to (4) and (6), for a given $k \geq 0$, yields

$$\hat{\theta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\bar{\mathbf{y}} \tag{16}$$

where

$$\text{Var}(\hat{\theta}) = \Sigma_{\hat{\theta}} = (\mathbf{X}^T\mathbf{X})^{-1}\sigma_e^2/r. \tag{17}$$

After fitting the $\theta$ effects the remaining components of error in (4) are the mean noise effects $\bar{\epsilon}$ and possible 'model error' due to the exclusion of significant higher order terms. In the manner of Draper and Smith [4, pp. 26-30] the model error $\mu_i$ at load $L_i$ is taken to be

$$\mu_i = \tau_i - \text{E}[\hat{p}_k(L_i)] \quad (i = 1, 2, \ldots, n)$$

where $\tau_i$ is the 'true' device response to the axial load $L_i$ and $\hat{p}_k(L_i)$ is as in (1). If the model is correct then $\mu_i = 0$ $(i = 1, 2, \ldots, n)$. The individual model error values $\{\mu_i\}$ cannot be estimated (since the true values $\{\tau_i\}$ are unknown), but their variance $\sigma_m^2 = \sum_{i=1}^n \mu_i^2/n$ can be estimated as will be shown later ($\bar{\mu} = \sum_{i=1}^n \mu_i/n$ is assumed to be zero).

After obtaining least squares estimates for $\theta$ by (16), under the $k$-degree model, the residual sum of squares in (4) is

$$\text{SS}_{em}(k) = (\bar{\mathbf{y}} - \mathbf{X}\hat{\theta})^T(\bar{\mathbf{y}} - \mathbf{X}\hat{\theta}) \tag{18}$$

so that

$$\text{E}[\text{SS}_{em}(k)] = (n - k - 1)(\sigma_e^2/r + \sigma_m^2).$$

The following algorithm is used to determine from the mean data the 'best' degree, $k^*$, of the polynomial $p_k(L_i)$ modeling the 'true' response curve. Statistical hypothesis testing at levels of significance $\alpha_m$ and $\alpha_k$ is used to determine whether, for a given $k$, the term $\theta_k L_i^k$ is needed. The $k_{max}$-degree model is fit first, then insignificant higher degree terms are dropped one by one. The term $\theta_k L_i^k$ is dropped if *either* $\theta_k$ *or* the model error under the $(k-1)$-degree model is determined to be not significant. The constraint $\max\{h^*, k_{min}\} \leq k^* \leq k_{max}$ is applied where $k_{min} = 1$. It is fully expected that for some force sensors the model error cannot be eliminated.

### Algorithm K.

1. Set $k = k_{max}$.

2. If $k \leq \max\{h^*, k_{min}\}$ then step 7. Else,

3. Set $\nu_{em} = n - k - 1$ and compute the statistic

$$F_m = \frac{r\text{SS}_{em}(k)/\nu_{em}}{\hat{\sigma}_e^2}$$

which has the $F(\nu_{em}, \nu_e)$ distribution under the hypothesis $\sigma_m^2 = 0$. Use (16) and (18).

13

4. If $G_{F(\nu_{em}, \nu_e)}(F_m) < 1 - \alpha_m$ or $k = k_{max}$ set $k = k - 1$ then step 2. Else,

5. Set $k = k + 1$, $\nu_{em} = n - k - 1$, and compute the statistic

$$F_k = \frac{\mathrm{SS}_{em}(k-1) - \mathrm{SS}_{em}(k)}{\mathrm{SS}_{em}(k)/\nu_{em}}$$

which has the $F(1, \nu_{em})$ distribution under the hypothesis $\theta_k = 0$.

6. If $G_{F(1,\nu_{em})}(F_k) < 1 - \alpha_k$ set $k = k - 2$ then step 2. Else,

7. Set $k^* = k$,

$$
\begin{aligned}
\nu_{em} &= n - k^* - 1, \quad \text{and} & (19) \\
\hat{\sigma}^2_{em} &= r\mathrm{SS}_{em}(k^*)/\nu_{em}. & (20)
\end{aligned}
$$

8. Done.

Having determined $h^*$ and $k^*$, the 'design' matrix $\mathbf{X}$ is fixed and $\hat{\boldsymbol{\theta}}$ is computed by (16). The estimated variance of $\hat{\boldsymbol{\theta}}$ is obtained by substituting the proper estimate of $\sigma^2_e$ into (17). When $\hat{\sigma}^2_{em}$ is *significantly larger* than $\hat{\sigma}^2_e$, indicated by a large value of $F_m$ in step 3 of algorithm $K$, then $\hat{\sigma}^2_{em}$ should be used so that the effects of significant model error are included in the estimated variance of $\hat{\boldsymbol{\theta}}$, thus

$$\hat{\boldsymbol{\Sigma}}_{\hat{\theta}} = (\mathbf{X}^T\mathbf{X})^{-1}\hat{\sigma}^2_{em}/r. \tag{21}$$

(Note that model errors are actually biases, and the quantity $\hat{\sigma}^2_{em}$ contains the mean squared model error as a reasonable estimate of bias effects.)

When $\hat{\sigma}^2_{em}$ is *not significantly larger* than $\hat{\sigma}^2_e$ the two values may be pooled to give

$$\hat{\sigma}^2_{ep} = (\nu_e\hat{\sigma}^2_e + \nu_{em}\hat{\sigma}^2_{em})/\nu_{ep}$$

where

$$\nu_{ep} = \nu_e + \nu_{em} = r(n - h^* - 1) + h^* - k^*, \tag{22}$$

thus

$$\hat{\boldsymbol{\Sigma}}_{\hat{\theta}} = (\mathbf{X}^T\mathbf{X})^{-1}\hat{\sigma}^2_{ep}/r. \tag{23}$$

Recalling from (5) that

$$\theta_t = \begin{cases} \beta_t + \bar{u}_t & (0 \le t \le h^*) \\ \beta_t & (h^* < t \le k^*) \end{cases} \tag{24}$$

and that $\mathrm{E}(\bar{u}_t) = 0$, an estimate of $\beta$ is

$$\hat{\beta} = \hat{\boldsymbol{\theta}} \tag{25}$$

with

$$\widehat{\mathrm{Var}}(\hat{\beta}) = \hat{\boldsymbol{\Sigma}}_{\hat{\beta}} = \hat{\boldsymbol{\Sigma}}_{\hat{\theta}} + \frac{1}{r}\begin{pmatrix} \hat{\boldsymbol{\Sigma}}_u & 0 \\ 0 & 0 \end{pmatrix} \tag{26}$$

where $\hat{\boldsymbol{\Sigma}}_{\hat{\theta}}$ is computed from either (21) or (23) as appropriate, $\hat{\boldsymbol{\Sigma}}_u$ is as in (14), and the zero matrices $(0)$ are of the proper size to make the rightmost matrix in (26) $(k^* + 1) \times (k^* + 1)$.

14

## 4.4   Summary of Error Contributions

To recapitulate from section 2.3, the five types of error considered are

1. roundoff,

2. within-run,

3. model,

4. between-run, and

5. systematic.

A review of past calibrations of force sensors at NBS has shown that any of these sources of error can be dominant. It would seem to be useful then, for a given calibration, to estimate the relative contribution of error from each of the sources. One measure of error from a given source is a '$\sigma$-like' value to represent the 'average' contribution of that source to a single observation. Estimates of such values can be computed from the data for the first four sources. For the fifth source an absolute limit ($E$) is used for combined effects of all 'systematic' errors as described in section 2.3. It is not a '$\sigma$-like' value.

In table 1 the sources of error are listed along with their '$\sigma$-like' estimators ($S$). The $S$-values are for rough comparison of sources of error only. They are not intended to be 'corrected' values of standard errors previously computed. (In the past some methods, which are valid under certain conditions, have been given for 'correcting' error variances for roundoff error. See Eisenhart [6].) Hopefully the $S$-values will aid the experimenter in understanding the characteristics of the force sensor-machine combination and perhaps indicate areas where improvements could be made.

Table 1: Estimators of known sources of error

| Error type | Error 'estimator' | Reference eqs. |
|---|---|---|
| roundoff | $S_r = \sqrt{\frac{1}{12n}\sum_{i=1}^{n}\Delta_i^2}$ | (27) |
| within-run | $S_e = \sqrt{\max\{0, \hat{\sigma}_e^2 - S_r^2\}}$ | (13,27,28) |
| model | $S_m = \sqrt{\max\{0, \hat{\sigma}_{em}^2 - \hat{\sigma}_e^2\}}$ | (13,20,28,29) |
| between-run | $S_b = \sqrt{(h^* + 1)\hat{\sigma}_b^2/n}$ | (15) |
| systematic | $E = $ combined limits | |

At this point the method of determining roundoff error must be addressed. Often data from a digital readout is inherently rounded. In recording the data the user may round even further. Traditionally (see Snedecor and Cochran [24, p. 81]) the rounding error, $\rho$, is treated as a uniformly distributed random variable in the interval $[-\Delta/2, \Delta/2]$ where $\Delta$ is the smallest possible difference between two unequal numbers. It follows that $\text{E}(\rho) = 0$ and

$\text{Var}(\rho) = \Delta^2/12$. For the sake of generality suppose that the rounding interval of the readout device is $\Delta_i$ at load $L_i$ and that $\rho_{ij}$ is the rounding error corresponding to the observation $y_{ij}$. Then the 'mean squared' rounding error is

$$S_r^2 = \frac{1}{12n} \sum_{i=1}^{n} \Delta_i^2. \tag{27}$$

In models (4) and (7) the $\{\epsilon_{ij}\}$ values were said to represent the 'noise' in the measurement system. In fact, each $\epsilon_{ij}$ also includes the roundoff error $\rho_{ij}$.

The quantities in (13) and (20) can be split into components

$$\hat{\sigma}_e^2 \quad \leftrightarrow \quad S_e^2 + S_r^2 \tag{28}$$

$$\hat{\sigma}_{em}^2 \quad \leftrightarrow \quad S_e^2 + S_r^2 + S_m^2 \tag{29}$$

to yield estimates of $S_m^2$ and $S_e^2$ free of roundoff error. The variance for between-run effects, $\hat{\sigma}_b^2$, in (15) reflects the 'average' variance per nonzero $u_{tj}$. Multiplying this by $(h^* + 1)/n$ yields an 'average' variance per observation, $S_b^2$.

## 4.5   Confidence and Prediction Intervals

The final step in the calibration process is to compute uncertainty bounds for the quantities of interest. In the present case those bounds are

1. a *confidence interval* for the 'true' sensor response at a given load,

2. a *confidence band* for the 'true' sensor response curve over its entire range, and

3. a *prediction interval* for a new, independent observation with the sensor randomly oriented in the same machine.

The bounds in 2 and 3 may further be used for inverse prediction as discussed in section 4.6. Seber [23, ch. 5] serves as a reference for the computations which follow. In each of the above cases formulas are given for two-sided $100(1 - \delta)\%$ bounds where $\delta$ is a probability in the exclusive range (0,1).

Note that one item missing from the above list is the vector of fixed effects $\beta$. In most regression applications a confidence interval for $\beta$ would be of interest, but here it is not. If such an interval were of interest it could be formed using the quantities defined in (25) and (26), but the question of the appropriate number of degrees of freedom to assign to the variance in (26) would be difficult. A frighteningly complicated approximation is given in [8].

Incorporating (24) the 'true' sensor response to the axial load $L$ can be expressed

$$
\begin{aligned}
p_{k^*}(L) &= \sum_{t=0}^{k^*} \beta_t L^t = \sum_{t=0}^{h^*} (\theta_t - \bar{u}_t) L^t + \sum_{t=h^*+1}^{k^*} \theta_t L^t \\
&= \sum_{t=0}^{k^*} \theta_t L^t - \sum_{t=0}^{h^*} \bar{u}_t L^t.
\end{aligned}
$$

16

Taking $E(\bar{u}_t) = 0$ the predicted sensor response is

$$\hat{p}_{k^*}(L) = \sum_{t=0}^{k^*} \hat{\theta}_t L^t$$

$$= \lambda^T \hat{\theta} \tag{30}$$

where $\lambda = (\lambda_1 \ \lambda_2)^T = (1 \ L \ \ldots \ L^{h^*} \ | \ L^{h^*+1} \ \ldots \ L^{k^*})^T$, and

$$\widehat{\mathrm{Var}}[\hat{p}_{k^*}(L)] = \lambda^T \hat{\Sigma}_{\hat{\theta}} \lambda + \lambda_1 \hat{\Sigma}_u \lambda_1 / r$$

$$= U_1 + U_2. \tag{31}$$

A *confidence interval for the 'true' response at load* $L$, $p_{k^*}(L)$, is then

$$\hat{p}_{k^*}(L) \pm t_{1-\delta/2}(\nu_1)\sqrt{\widehat{\mathrm{Var}}[\hat{p}_{k^*}(L)]}$$

where $\nu_1$ is the number of degrees of freedom associated with the estimated variance and $t_{1-\delta/2}(\nu_1)$ is a quantile of Student's $t$ distribution with $\nu_1$ degrees of freedom (see appendix A). Since the estimated variance in (31) is a linear combination of two variances, an 'effective' number of degrees of freedom is computed by the Welch-Satterthwaite formula (see Gaylor and Hopper [7]) to be

$$\nu_1 = \frac{(U_1 + U_2)^2}{U_1^2/\nu_\theta + U_2^2/(r-1)} \tag{32}$$

where

$$\nu_\theta = \begin{cases} \nu_{em}, & \text{as in (19), if } \hat{\Sigma}_{\hat{\theta}} \text{ was computed by (21)} \\ \nu_{ep}, & \text{as in (22), if } \hat{\Sigma}_{\hat{\theta}} \text{ was computed by (23)} \end{cases}.$$

Note that if $h^* = -1$ then $\hat{\Sigma}_u$ is degenerate, thus $U_2 = 0$ and $\nu_1 = \nu_\theta$. This type of confidence interval is appropriate, for example, if the sensor response at nominal maximum load is to be used as a check standard. This type of interval is illustrated by the dotted lines in figure 12 for the example in appendix B.

A *confidence band for the entire 'true' response curve* is taken to be the set (over $L$) of all individual confidence intervals for $p_{k^*}(L)$ of the form

$$\hat{p}_{k^*}(L) \pm K_1 \tag{33}$$

where

$$K_1 = \sqrt{(k^* + 1)F_{1-\delta}(k^* + 1, \nu_1)\widehat{\mathrm{Var}}[\hat{p}_{k^*}(L)]} \tag{34}$$

and $F_{1-\delta}(k^* + 1, \nu_1)$ is a quantile of the $F$ distribution with $k^* + 1$ and $\nu_1$ degrees of freedom as described in appendix A. This type of band is illustrated by the dashed line in figure 12.

A third quantity of interest is a confidence interval for a new observation $y^*$ obtained when the sensor is re-oriented in the same measuring machine and a load $L^*$ applied. The new observation can be written

$$y^* = p_{k^*}(L^*) + \lambda_1^{*T}\mathbf{u} + \epsilon^*$$

17

where $\lambda^* = (\lambda_1^* \; \lambda_2^*) = \left(1 \; L^* \; \ldots \; L^{*h^*} \; | \; L^{*(h^*+1)} \; \ldots \; L^{*k^*}\right)^T$, $\mathbf{u}$ is a random $(h^* + 1)$-vector from a distribution with mean zero and variance $\Sigma_u$, and $\epsilon^*$ is an independent deviate from a distribution with mean zero and variance $\sigma_e^2$. It follows that

$$
\begin{aligned}
\mathrm{E}(y^*) &= p_{k^*}(L^*) \quad \text{and} & (35) \\
\mathrm{Var}(y^*) &= \mathrm{Var}[p_{k^*}(L^*)] + \lambda_1^{*T} \Sigma_u \lambda_1^* + \sigma_e^2 \\
&= \lambda_1^{*T} \Sigma_u \lambda_1^* + \sigma_e^2. & (36)
\end{aligned}
$$

Replacing the expectation and variance of $y^*$ by their estimates and incorporating (30) and (31) gives

$$
\begin{aligned}
\hat{y}^* &= \hat{p}_{k^*}(L^*) = \lambda^{*T}\hat{\theta} \quad \text{and} \\
\widehat{\mathrm{Var}}(\hat{y}^*) &= \widehat{\mathrm{Var}}[\hat{p}_{k^*}(L^*)] + \lambda_1^{*T} \hat{\Sigma}_u \lambda_1^* + \hat{\sigma}_e^2 \\
&= \lambda^{*T} \hat{\Sigma}_{\hat{\theta}} \lambda^* + (1 + \frac{1}{r})\lambda_1^{*T} \hat{\Sigma}_u \lambda_1^* + \hat{\sigma}_e^2 & (37) \\
&= V_1 + V_2 + V_3.
\end{aligned}
$$

As before the 'effective' number of degrees of freedom for the estimated variance in (37) is

$$
\nu_2^* = \frac{(V_1 + V_2 + V_3)^2}{V_1^2/\nu_\theta + V_2^2/(r-1) + V_3^2/\nu_e} \tag{38}
$$

where $\nu_\theta$ is as in (32) and $\nu_e$ is as in (12). As before, if $h^* = -1$ then $\hat{\Sigma}_u$ is degenerate and (38) is computed with $V_2 = 0$. A *confidence interval for a new observation $y^*$ as defined above* is then

$$
\hat{p}_{k^*}(L^*) \pm K_2 \tag{39}
$$

where

$$
K_2 = t_{1-\delta/2}(\nu_2^*)\sqrt{\widehat{\mathrm{Var}}(\hat{y}^*)}. \tag{40}
$$

This type of interval is illustrated by the solid line in figure 12.

## 4.6   Inverse Prediction

The calibrated force sensor is ordinarily used to measure an unknown load $L^*$ from a single new observation $y^*$ on some machine, not necessarily the one on which it was calibrated. The desired result is an estimate of the magnitude of the unknown load and a confidence interval for it. This application, known as *inverse prediction*, consists of

- computing the calibration curve and the appropriate confidence band(s) for the sensor,

- obtaining one or more new observations at unknown load(s), and

- estimating each unknown load and its confidence interval.

18

MULTIPLE USE INVERSE PREDICTION ON SAME MACHINE

eq. (45) →

RESPONSE

y*

LOAD
Figure 10

THREE TYPES OF 95% CONFIDENCE INTERVALS (AND ORIGINAL DATA)

NUMERALS DESIGNATE DATA AND RUN NUMBER

DOTTED = C.I. FOR SINGLE POINT ON CURVE
DASHED = C.I. FOR ENTIRE CURVE
SOLID = C.I. FOR A NEW OBSERVATION

DEVIATION FROM FITTED CURVE (×10**5)

LOAD (LbF)
Figure 12

SINGLE USE INVERSE PREDICTION ON SAME MACHINE

eq. (39) →

RESPONSE

y*

LOAD
Figure 9
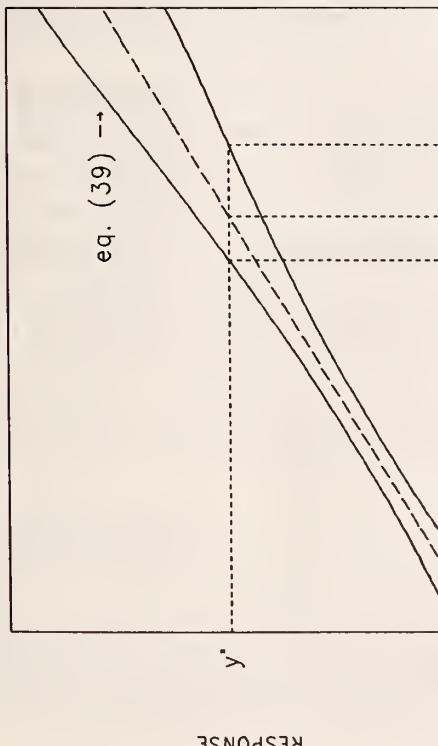
MULTIPLE USE INVERSE PREDICTION ON DIFFERENT MACHINE

eq. (33) →

RESPONSE

$y^* + K_4$
$y^*$
$y^* - K_4$

LOAD
Figure 11

19
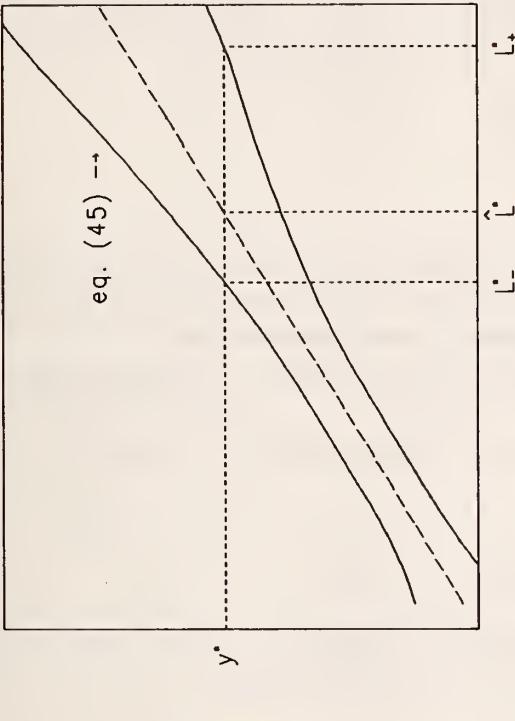
Three cases of inverse prediction with a calibrated force sensor will be considered here:

- a *single use* case on the same machine,

- a *multiple use* case on the same machine, and

- a *multiple use* case on a different machine.

The term *multiple use* simply means that the same calibration curve will be used for inverse prediction more than once. This situation is believed most likely to occur in practice, thus the *multiple use* methods (described later in this section) are recommended.

The *single use* case will be given only for use on the same machine. A confidence interval for a new observation $y^*$ can be obtained for any $L$ by (39). So, given $y^*$, a $100(1 - \delta)\%$ confidence interval for the unknown load $L^*$ is the set of $L$'s for which the confidence interval calculated by (39) encloses $y^*$, i.e., for which

$$\hat{p}_{k^*}(L) - K_2 \leq y^* \leq \hat{p}_{k^*}(L) + K_2$$

where $K_2$ is as in (40). This set will be connected when the upper and lower bounds of the confidence band are monotonic. The estimate of $L^*$ is $\hat{L}^*$ which satistifes

$$\hat{p}_{k^*}(\hat{L}^*) = y^*.$$

If $k^* > 1$ then $\hat{L}^*$ may have to be computed iteratively. Figure 9 illustrates the graphical procedure for determining $\hat{L}^*$ and the confidence interval $(L^*_-, L^*_+)$ for $L^*$. Again, this type of confidence interval for $L^*$ is valid only when a *single* inverse prediction is made after a given calibration.

In the *multiple use* case (*same* machine) a well known but complicated method of inverse prediction is that of Scheffé [20]. In a forthcoming paper Carroll, Sacks, and Spiegelman [3], hereafter designated *CSS*, present a modification of the Scheffé method which is far simpler to implement and generally leads to shorter intervals. The first part of the *CSS* method is to compute the $100(1 - \delta)\%$ confidence band for the entire response curve as in (33) and (34). This band is illustrated in figure 11. The second part is to compute a $100(1 - \alpha)\%$ confidence interval for the 'true' response $p_{k^*}(L^*)$ based on the new observation $y^*$. The estimated variance of $y^*$ is computed from (36) to be

$$\begin{aligned} \widehat{\mathrm{Var}}(y^*) &= \lambda_1^{*T} \hat{\Sigma}_u \lambda_1^* + \hat{\sigma}_e^2 \\ &= W_1 + W_2 \end{aligned} \tag{41}$$

which is 'effectively' based on

$$\nu_3^* = \frac{(W_1 + W_2)^2}{W_1^2/(r - 1) + W_2^2/\nu_e} \tag{42}$$

20

degrees of freedom. If $\hat{\Sigma}_u$ is degenerate then $W_1 = 0$ and $\nu_3^* = \nu_e$. Notice that the variance of $\hat{p}_{k^\bullet}(L^*)$ does not enter into this part of the computation. An appropriate confidence interval for $p_{k^\bullet}(L^*)$, incorporating (35) and (41), is

$$y^* \pm K_3 \qquad (43)$$

where

$$K_3 = t_{1-\alpha/2}(\nu_3^*)\sqrt{\widehat{\mathrm{Var}}(y^*)} \qquad (44)$$

Note that the variance term $W_1$ in (41) is a function of $L^*$ (through $\lambda_1^*$) which is unknown. In this case $L^*$ may be approximated by $\hat{p}_{k^\bullet}^{-1}(y^*)$ in the vector $\lambda_1^*$.

A confidence interval for the unknown load $L^*$, using the *CSS* procedure, is then the set of all $L$'s for which

$$\hat{p}_{k^\bullet}(L) - K_1 - K_3 < y^* < \hat{p}_{k^\bullet}(L) + K_1 + K_3 \qquad (45)$$

where $K_1$ and $K_3$ are as in (34) and (44) respectively. This procedure assures that, for a given calibration process, the probability is $1 - \delta$ that the *expected* proportion of calculated confidence intervals for $L^*$ that actually cover $L^*$ will be at least $1 - \alpha$ in the long run. The probability $1 - \delta$ refers to the probability that the calibration of the sensor is 'good', that is, the confidence band covers the 'true' underlying response curve. Figure 10 illustrates the graphical procedure for determining $\hat{L}^*$ and the confidence interval $(L_-^*, L_+^*)$ for $L^*$ in the case under consideration.

In the *multiple use* case (*different* machine from the one on which the sensor was calibrated) the previous analysis may be applied using the variance estimates $\hat{\Sigma}_u$ and $\hat{\sigma}_e^2$ appropriate to that machine. If these quantities are available then they, along with their degrees of freedom which correspond to $r - 1$ and $\nu_e$, may be substituted into (41) and (42) to give $K_3$ as in (44). The prediction of $L^*$ and its confidence band then follow as before. In the likely event that both of these quantities are *not* available, an approximate $100(1 - \alpha)\%$ confidence interval for $p_{k^\bullet}(L^*)$ of the form

$$y^* \pm K_4$$

corresponding to (43) may be used where $K_4$ is based on statistical evidence or 'engineering judgement'. This interval is used in conjunction with (33) to provide an estimate $\hat{L}^*$ and a confidence interval $(L_-^*, L_+^*)$ for $L^*$ as illustrated in figure 11. The probability statement for the coverage is the same as in the previous case (fig. 10).

In *CSS* two modifications are discussed which will shorten the inverse prediction intervals slightly, but for the sake of simplicity they will not be used here.

# 5 Example

In the process of developing the model described in this paper, many sets of force calibration data were analyzed. The complete analysis of one of these data sets will now be illustrated. The example chosen is typical in that

- three runs were made,

- the between-run error appears to be well modeled by linear pieces, and

- the residuals from a quadratic fit appear to have a serpentine form.

The example was analyzed by the new method with $h_{max} = 3$, $k_{max} = 5$, $\alpha_h = 0.01$, $\alpha_m = 0.10$, and $\alpha_k = 0.025$. At each stage of the fitting procedure to determine $h^*$ the probability of overfitting was $\alpha_h = 0.01$, and the corresponding probability of overfitting when determining $k^*$ was $\alpha_m \alpha_k = 0.0025$. The analysis was performed by the FORTRAN program FCAL88 which incorporates the new method described in this paper. The output from the program is given in appendix B. The net observations $\{y_{ij}\}$ appear on the first page of that output. The degrees of the 'best fitting' polynomials were determined to be $h^* = 1$ and $k^* = 3$. Figures 1 through 12 are based on this analysis of the example. In figures 9 through 11 the confidence intervals have been magnified by a factor of 2000 vertically for the purpose of illustration.

The same observations were re-run with $h_{max} = k_{max} = 2$ in order to limit the fits to nothing higher than a quadratic. In that case $h^* = 1$ and $k^* = 2$ were computed (the computer output is not shown). Some results from the two analyses are compared in table 2 below.

Table 2: Comparison of two fits to the example data

| Quantity | k*=2 | k*=3 | Reference |
|---|---|---|---|
| $h^*$ | 1 | 1 | (Alg. H) |
| $\hat{\sigma}_e$ | $0.120 \cdot 10^{-4}$ | $0.120 \cdot 10^{-4}$ | Eq. (13) |
| $\hat{\sigma}_{em}$ | $0.754 \cdot 10^{-4}$ | $0.204 \cdot 10^{-4}$ | Eq. (20) |
| $\hat{\theta}_0$ | 1.03347 | 1.03347 | Eq. (16) |
| $\hat{\theta}_1$ | 0.96743 | 0.96767 | Eq. (16) |
| $\hat{\theta}_2$ | $0.32926 \cdot 10^{-3}$ | $0.32575 \cdot 10^{-3}$ | Eq. (16) |
| $\hat{\theta}_3$ | — | $-0.33087 \cdot 10^{-3}$ | Eq. (16) |
| $S_r$ | $0.2887 \cdot 10^{-5}$ | $0.2887 \cdot 10^{-5}$ | (Table 1) |
| $S_e$ | $0.1167 \cdot 10^{-4}$ | $0.1167 \cdot 10^{-4}$ | (Table 1) |
| $S_m$ | $0.1301 \cdot 10^{-3}$ | $0.3329 \cdot 10^{-4}$ | (Table 1) |
| $S_b$ | $0.2263 \cdot 10^{-4}$ | $0.2263 \cdot 10^{-4}$ | (Table 1) |
| $E$ | $0.2013 \cdot 10^{-4}$ | $0.2013 \cdot 10^{-4}$ | (Table 1) |

Note that FCAL88 produces a two-page table with four kinds of 95% confidence/prediction intervals. This allows inverse prediction to be done by interpolation on the appropriate columns. For example, suppose an unknown load $L^*$ was applied (on the same machine) and that an independent (net) observation $y^* = 1.47901$ was obtained. The estimate of $L^*$ linearly interpolated from the **Predicted response** and **Load** columns is $L^* = 221,757.2$

(see table 3 which reproduces part of the two-page table from appendix B). Using columns

Table 3: Partial listing of 95% confidence intervals for the example data

| Load | Predicted response | | Single use | | | Multiple use | |
| | | | [3] (lower) | [4] (upper) | | [7] (lower) | [8] (upper) |
|---|---|---|---|---|---|---|---|
| 219,000 | 1.46061 | ... | 1.46051 | 1.46071 | ... | 1.46044 | 1.46078 |
| 222,000 | 1.48063 | ... | 1.48053 | 1.48074 | ... | 1.48046 | 1.48081 |

[3] and [4] the *single use* 95% confidence interval for $L^*$ is (221,740.9 , 221,772.2), and using columns [7] and [8] the corresponding *multiple use* interval is (221,730.4 , 221,782.7). If uncertainty due to systematic errors is to be accounted for, each of these confidence intervals would be increased by $(0.002\%)(221,757.2) \approx 4.4$ in each direction. (For this example the load unit is pounds-force.)

For numerical stability FCAL88 linearly transforms the load values $L_i$ to the interval $[-1, 1]$ before any computations begin (a good practice in polynomial regression). This transformation is invisible to the user except for the computed values of $\hat{\Sigma}_u$ ('ESTIMATED COVARIANCE MATRIX FOR BETWEEN-RUN EFFECTS', p. B-3), $\hat{\theta}$ ('ESTIMATES OF POLYNOMIAL RESPONSE CURVE COEFFICIENTS FROM MEAN DATA', p. B-7), and $\hat{\Sigma}_{\hat{\theta}}$ ('ESTIMATED COVARIANCE MATRIX FOR POLYNOMIAL COEFFICIENTS', p. B-7). These three quantities are highly dependent on the extreme values of the load grid,

$$L_{min} = \min\{L_1, L_2, \ldots, L_n\} \quad \text{and} \quad L_{max} = \max\{L_1, L_2, \ldots, L_n\} \ ,$$

which transform to $-1$ and $1$ respectively. Consequently, values of $\hat{\Sigma}_u$, $\hat{\theta}$, and $\hat{\Sigma}_{\hat{\theta}}$ obtained from different calibrations are comparable *only if* $L_{min}$ and $L_{max}$ are the same in each case.

The linear transformation to $[-1, 1]$ is

$$x = \frac{2(L - L_{min})}{L_{max} - L_{min}} - 1 \ , \tag{46}$$

and the inverse transformation is

$$L = \frac{(x + 1)(L_{max} - L_{min})}{2} + L_{min} \ . \tag{47}$$

The original and transformed load values for the example data are shown in table 4. The virtue in transforming is seen when examining, for example, $\hat{\Sigma}_{\hat{\theta}}$ as defined by (21) or (23). This quantity contains the factor

$$\mathbf{U} = (\mathbf{X}^T \mathbf{X})^{-1} \tag{48}$$

which is called the *unscaled* variance of $\hat{\theta}$. Note that $\mathbf{U}$ depends only on the pre-selected load grid, by (3), and not on the observed data, therefore it is an appropriate means for

Table 4: Original and transformed load values for the example data

| $i$ | Original $L_i$ (lbf) | Transformed $x_i$ |
|---|---|---|
| 1 | 10,000 | −1.000000 |
| 2 | 30,000 | −0.862069 |
| 3 | 60,000 | −0.655172 |
| 4 | 90,000 | −0.448276 |
| 5 | 120,000 | −0.241379 |
| 6 | 150,000 | −0.034483 |
| 7 | 180,000 | +0.172414 |
| 8 | 210,000 | +0.379310 |
| 9 | 240,000 | +0.586207 |
| 10 | 270,000 | +0.793103 |
| 11 | 300,000 | +1.000000 |

comparison. Using the $L_i$ values to form $\mathbf{X}$, equation (48) yields

$$
\mathbf{U}_L = \begin{pmatrix}
0.116 \cdot 10^{+01} & -0.295 \cdot 10^{-04} & 0.194 \cdot 10^{-09} & -0.368 \cdot 10^{-15} \\
-0.295 \cdot 10^{-04} & 0.104 \cdot 10^{-08} & -0.772 \cdot 10^{-14} & 0.156 \cdot 10^{-19} \\
0.194 \cdot 10^{-09} & -0.772 \cdot 10^{-14} & 0.610 \cdot 10^{-19} & -0.128 \cdot 10^{-24} \\
-0.368 \cdot 10^{-15} & 0.156 \cdot 10^{-19} & -0.128 \cdot 10^{-24} & 0.277 \cdot 10^{-30}
\end{pmatrix}. \tag{49}
$$

Similarly, using the corresponding $x_i$ values to form $\mathbf{X}$,

$$
\mathbf{U}_x = \begin{pmatrix}
0.214 \cdot 10^{+00} & -0.149 \cdot 10^{-02} & -0.296 \cdot 10^{-00} & -0.164 \cdot 10^{-02} \\
-0.149 \cdot 10^{-02} & 0.157 \cdot 10^{+01} & 0.160 \cdot 10^{-02} & -0.186 \cdot 10^{+01} \\
-0.296 \cdot 10^{+00} & 0.160 \cdot 10^{-02} & 0.712 \cdot 10^{+00} & 0.272 \cdot 10^{-01} \\
-0.164 \cdot 10^{-02} & -0.186 \cdot 10^{+01} & 0.272 \cdot 10^{-01} & 0.257 \cdot 10^{+01}
\end{pmatrix}. \tag{50}
$$

Note that the lower right entry in the right hand side of (49) approaches the underflow limit on some computers, a potentially unpleasant situation. A much better situation occurs when computing with the transformed loads, as illustrated in (50). Transitions to and from the original loads are easily made with (46) and (47).

# 6  Simulation Results

A simulation study was performed using the estimated quantities computed for the example (by the new method) as a basis. Specifically, 10,000 sets of $\{y_{ij}\}$ values were generated using (2) where $(\hat{\beta}_0\ \hat{\beta}_1\ \hat{\beta}_2\ \hat{\beta}_3)$ were substituted for $(\beta_0\ \beta_1\ \beta_2\ \beta_3)$, the $\{\epsilon_{ij}\}$ were generated from a normal distribution with mean zero and variance $\hat{\sigma}_e^2$, and the $\{(u_{0j}\ u_{1j})\}$ were generated from a multivariate normal distribution with mean zero and variance $\hat{\Sigma}_u$. No allowance was made for model error. Each set of simulated observations was run under the FCAL88 program. The number of occurrences of each computed $(h^*, k^*)$ pair are given in table 5

24

Table 5: Computed 'best' polynomial degrees for 10,000 simulated data sets

| | $k^*$ | | | | |
|---|---|---|---|---|---|
| $h^*$ | 1 | 2 | 3 | 4 | 5 |
| $-1$ | 0 | 0 | 1683 | *2* | *0* |
| 0 | 0 | 0 | 1118 | *1* | *0* |
| 1 | 0 | 0 | 6949 | *23* | *17* |
| 2 | *0* | *0* | *104* | *1* | *0* |
| 3 | *0* | *0* | *99* | *0* | *3* |

above. The levels of significance used were $\alpha_h = 0.01$, $\alpha_m = 0.10$, and $\alpha_k = 0.025$ as before. Note that the entries in *italics* represent cases of overfitting which are controlled by the three levels of significance. For example, the number of entries in the $h^* = 3$ row is 102 which approximates the expected number $10,000\alpha_h = 100$, and the number of entries in the $k^* = 5$ column is 20 which approximates the expected number $10,000\alpha_m\alpha_k = 25$. Similar numbers of entries would be expected for the cases $h^* = 2$ and $k^* = 4$ respectively.

Underfitting is little cause for concern in this simulation. The fact that no cases of $k^* = 1$ or 2 were observed indicates that the coefficient $\beta_3$ is highly significant in the example. On the other hand, the coefficients of the linear terms in the between-run effects are much less significant since they were determined to be not significant about 28% of the time.

The greatest value in this simulation exercise may be to forewarn the user that repeated calibrations of the same force sensor may result in several different $(h^*, k^*)$ pairs, even though the underlying 'true' sensor response curve does not change.

# 7   Optimality Considerations

In a typical NBS calibration of a force sensor the applied loads are uniformly spaced between 10% and 100% of sensor capacity (see section 2.2). The question naturally arises as to whether a different grid of applied loads, with the same number of points, would produce 'better' results in some sense. Seber [23, pp. 231-234] discusses several types of optimality criteria for polynomial regression of a fixed degree. Of those criteria the one of most importance in the present application would seem to be the minimization of the variance of $\hat{\theta}_k$, the fitted coefficient of the highest power of the load. In algorithm $K$, $\hat{\theta}_k$ is tested for significance for one or more values of $k$.

For a given load grid and polynomial degree $k$ the 'design' matrix $\mathbf{X}$ is computed as in (3). This matrix is then used in computing the vector of fitted coefficients $\hat{\theta}$ as in (16) and its variance as in (21) or (23). As discussed in section 5, the factor $(\mathbf{X}^T\mathbf{X})^{-1}$ is the *unscaled* variance of $\hat{\theta}$, thus the unscaled variance of $\hat{\theta}_k$ is the $(k+1, k+1)$ entry of $(\mathbf{X}^T\mathbf{X})^{-1}$ (lower right corner). This quantity is computed for polynomial fits of degree 2, 3, 4, and 5 for each of seven different load grids as shown in table 6. (In this paper it is assumed that no

Table 6: Comparison of unscaled variances of the fitted coefficients $\hat{\theta}_k$ for several load grids with $n = 10$

| Load grid (% of sensor capacity) | Unscaled variance of $\hat{\theta}_k$ | | | |
| --- | --- | --- | --- | --- |
| | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ |
| [1] 10.0 10.0 10.0 55.0 55.0<br>55.0 55.0 100.0 100.0 100.0 | **0.417** | — | — | — |
| [2] 10.0 10.0 32.5 32.5 32.5<br>77.5 77.5 77.5 100.0 100.0 | 0.741 | **1.63** | — | — |
| [3] 10.0 10.0 23.2 23.2 55.0<br>55.0 86.8 86.8 100.0 100.0 | 0.714 | 3.00 | **7.00** | — |
| [4] 10.0 18.6 18.6 41.1 41.1<br>68.9 68.9 91.4 91.4 100.0 | 0.800 | 3.20 | 12.79 | **25.6** |
| [5] 10.0 12.7 20.5 32.5 47.2<br>62.8 77.5 89.5 97.3 100.0 | 0.740 | 3.01 | 12.19 | 49.3 |
| [6] 10.0 15.0 20.0 30.0 45.0<br>65.0 80.0 90.0 95.0 100.0 | 0.874 | 3.11 | 11.63 | 40.4 |
| [7] 10.0 20.0 30.0 40.0 50.0<br>60.0 70.0 80.0 90.0 100.0 | 0.777 | 2.69 | 10.21 | 43.7 |

[1] Optimal load grid in 10%–100% range for $2^{nd}$ degree fit
[2] Optimal load grid in 10%–100% range for $3^{rd}$ degree fit
[3] Optimal load grid in 10%–100% range for $4^{th}$ degree fit
[4] Optimal load grid in 10%–100% range for $5^{th}$ degree fit
[5] Optimal load grid in 10%–100% range for $9^{th}$ degree fit
[6] Load grid [5] rounded to nearest 5%
[7] Uniform load grid

polynomial higher than a $5^{th}$ degree will be fit.) In order for the comparisons to be fair, each grid contains the same number of points. The first five of these grids are optimal for the indicated polynomial degrees. They are the so-called Chebyshev points,

$$x_i = -\cos\left(\frac{i\pi}{k}\right) \quad (i = 0, 1, \ldots, k),$$

linearly transformed from the interval $[-1, 1]$ to the interval $[10\%, 100\%]$ times sensor capacity. Note that in the case $n > k + 1$ optimality is achieved by replicating measurements at certain loads. In all cases the unscaled variance of $\hat{\theta}_k$ was computed from the equivalent loads in $[-1, 1]$.

The first four grids provide the optimal values shown in boldface. The first three are of no use since there are not enough distinct loads to fit all polynomials of interest. Examination of the last four shows that none is consistently better than the others over the degrees

considered. By a slim margin the uniform grid appears to be best. Based on this brief analysis there seems to be no evidence that any other load grid should be preferred over the uniform.

# 8   Conclusion

Force sensors are not characterized by a single number but by a response curve over a wide range of loads. The precise form of this curve depends on the details of design and construction of the particular sensor. After trying various mathematical models for the response curve, it was concluded that variable degree polynomials generally provided the best fits.

A new statistical model for calibrating force sensors, based on automatic procedures for determining the degrees of polynomials, has been presented. It attempts to model several known sources of error and quantize their relative contributions. Methods for computing several types of confidence and prediction intervals have been given. The three levels of significance used in determining the polynomial degrees may be adjusted to suit the users needs as may be the levels for the confidence/prediction intervals. Usage of the calibration curves for inverse prediction has also been discussed and illustrated. A brief optimality study has shown that the traditional uniform spacing of applied loads cannot be improved on when allowing polynomial fits of $5^{th}$ degree or less.

This new method of analysis is not likely to solve all the problems of force sensor calibration. The author believes, however, that this method will prove to be a useful tool in developing a better understanding and a better characterization of their behavior.

---

# References

[1] American Society for Testing and Materials, 'Standard Practice of Calibration of Force-Measuring Instruments for Verifying the Load Indication of Testing Machines', E74-83, *Annual Book of ASTM Standards*, 1984.

[2] Cameron, J.M., 'The Use of the Method of Least Squares in Calibration', National Bureau of Standards, *NBSIR 74-587*, September 1974.

[3] Carroll, R.J., Sacks, J., and Spiegelman, C.H., 'A Quick and Easy Multiple Calibration Curve Procedure', to appear in *Technometrics*, May 1988.

[4] Draper, N.R. and Smith, H., *Applied Regression Analysis*, John Wiley & Sons, Inc., 1966.

[5] Cunningham, E.P. and Henderson, C.R., 'An Iterative Procedure for Estimating Fixed Effects and Variance Components in Mixed Model Situations', *Biometrics 24(1)*, March 1968, pp. 13-25.

[6] Eisenhart, C., 'Effects of Rounding or Grouping Data', Ch. 4, pp. 186-223 in Eisenhart, C., Hastay, M.W., and Wallis, W.A., *Selected Techniques for Statistical Analysis*, McGraw-Hill Book Company, 1947.

[7] Gaylor, D.W. and Hopper, F.N., 'Estimating the Degrees of Freedom for Linear Combinations of Mean Squares by Satterthwaite's Formula', *Technometrics 11(4)*, November 1969, pp. 691-706.

[8] Geisbrecht, F.G. and Burns, J.C., 'Two-Stage Analysis on a Mixed Model: Large-Sample Asymptotic Theory and Small-Sample Simulation Results', *Biometrics 41(2)*, June 1985, pp. 477-486.

[9] Hager, H. and Antle, C., 'The Choice of Degree of a Polynomial Model', *Journal of the Royal Statistical Society B 30(3)*, 1968, pp. 469-471.

[10] Henderson, Charles R., 'Analysis of Covariance in the Mixed Model: Higher-Level, Nonhomogeneous, and Random Regressions', *Biometrics 38(3)*, September 1982, pp. 623-640.

[11] Hockersmith, Thomas E. and Ku, Harry H., 'Uncertainties Associated with Proving Ring Calibration', *Precision Measurement and Calibration – Mechanics*, National Bureau of Standards, *SP 300, Vol. 8*, pp. 239-246 (also, Instrument Society of America, *Preprint No. 12.3-2-64*, 19th Annual ISA Conference and Exhibit, October 1964).

[12] Hoel, Paul G., 'On Testing for the Degree of a Polynomial', *Technometrics 10(4)*, November 1968, pp. 757-767.

[13] Mitchell, Richard A., 'Force Calibration at the National Bureau of Standards', National Bureau of Standards, *Technical Note 1227*, August 1986.

[14] Mitchell, R.A. and Baker, S.M., 'Characterizing the Creep Response of Load Cells', *VDI Berichte, Nr. 312*, 1978, pp. 43-48.

[15] Mitchell, Richard A. and Pontius, Paul E., 'Force Sensor-Machine Interaction', *Proceedings of the 27th International Instrumentation Symposium*, Instrument Society of America, 1981, pp. 225-232.

[16] Mitchell, R.A., Seifarth, R.L., and Reeve, C.P., 'Eccentric Loading Sensitivity of Force Sensors', *Proceedings of the 11th Conference of IMEKO TC-3 on Measurement of Force and Mass*, Amsterdam, The Netherlands, May 1986.

[17] Peterson, R.W. and Bloss, R.L., 'Interlaboratory Comparison of Force Calibration Using ASTM Method E74-74, Phase I', National Bureau of Standards, *NBSIR 76-1145*, August 1976.

[18] Peterson, Robert W., Jenkins, Linwood, and Mitchell, Richard A., 'Interlaboratory Comparison of Force Calibrations Using ASTM Method E74-74', National Bureau of Standards, *Technical Note 1211*, April 1985.

[19] Pontius, P.E. and Mitchell, R.A., 'Inherent Problems in Force Measurement', *Experimental Mechanics 22(3)*, March 1982, pp. 81-88.

[20] Scheffé, H., 'A Statistical Theory of Calibration', *Annals of Statistics 1(1)*, January 1973, pp. 1-37.

[21] Searle, S.R., 'Topics in Variance Component Estimation', *Biometrics 27(1)*, March 1971, pp. 1-76.

[22] Searle, S.R., *Linear Models*, John Wiley & Sons, 1971.

[23] Seber, G.A.F., *Linear Regression Analysis*, John Wiley & Sons, 1977.

[24] Snedecor, George W. and Cochran, William G., *Statistical Methods*, Sixth Edition, The Iowa State University Press, 1976.

[25] Wilson, Bruce L., Tate, Douglas R., and Borkowski, George, 'Proving Rings for Calibrating Testing Machines', National Bureau of Standards, *Circular C454*, August 1946.

# A    Definitions of the F and t distributions

The *probability density function (p.d.f.)* of the $F(m, n)$ distribution may be expressed as

$$g_{F(m,n)}(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m-2}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\left(1 + \frac{mx}{n}\right)^{\frac{m+n}{2}}}$$

where $x \geq 0$, $m > 0$, $n > 0$, and $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$. This function is plotted in figure 13 for the case $m = 5$ and $n = 18$.

The *cumulative distribution function (c.d.f.)* of the $F(m, n)$ distribution may be expressed as

$$G_{F(m,n)}(x) = \int_0^x g_{F(m,n)}(u)du$$

where $x \geq 0$. This function is plotted in figure 14 for $m$ and $n$ as above. Note that $G_{F(m,n)}(0) = 0$ and $G_{F(m,n)}(\infty) = 1$.

The *p-quantile* of the $F(m, n)$ distribution, denoted by $F_p(m, n)$, is the $x$ value which satisfies $G_{F(m,n)}(x) = p$.

When an $F$ test is performed either the *c.d.f.* or a *quantile* must be computed. Since the c.d.f. is easier to compute on a machine, that quantity has been incorporated into the $F$ tests of algorithms $H$ (section 4.2) and $K$ (section 4.3).

In similar fashion the *p.d.f.* of the $t(n)$ distribution is

$$g_{t(n)}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{\pi n}\ \Gamma\left(\frac{n}{2}\right)}$$

where $-\infty < x < \infty$ and $n > 0$. This curve is plotted in figure 15 for the case n=10. The *c.d.f.* is

$$G_{t(n)}(x) = \int_{-\infty}^x g_{t(n)}(u)du$$

where $-\infty < x < \infty$. Note that $G_{t(n)}(-\infty) = 0$, $G_{t(n)}(0) = 0.5$, and $G_{t(n)}(\infty) = 1$. This function is plotted in figure 16 for n as above. The *p-quantile*, denoted by $t_p(n)$, is the $x$ value which satisfies $G_{t(n)}(x) = p$.

In this paper the $t$ distribution is used only in computing confidence intervals (sections 4.5 and 4.6), thus only quantiles are required.

PROBABILITY DENSITY FUNCTION OF F(5,18) DISTRIBUTION

g(x)

X
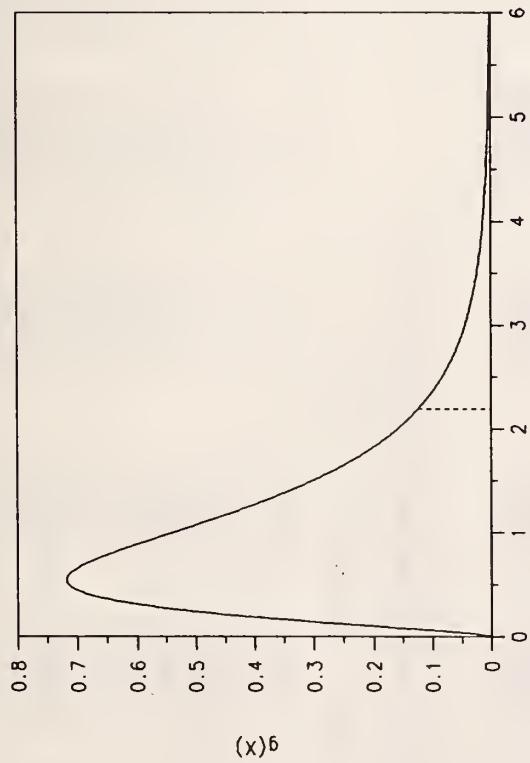
Figure 13

CUMULATIVE DISTRIBUTION FUNCTION OF F(5,18) DISTRIBUTION

$F_{0.9}(5,18) = 2.196$

G(x)

0.900

X
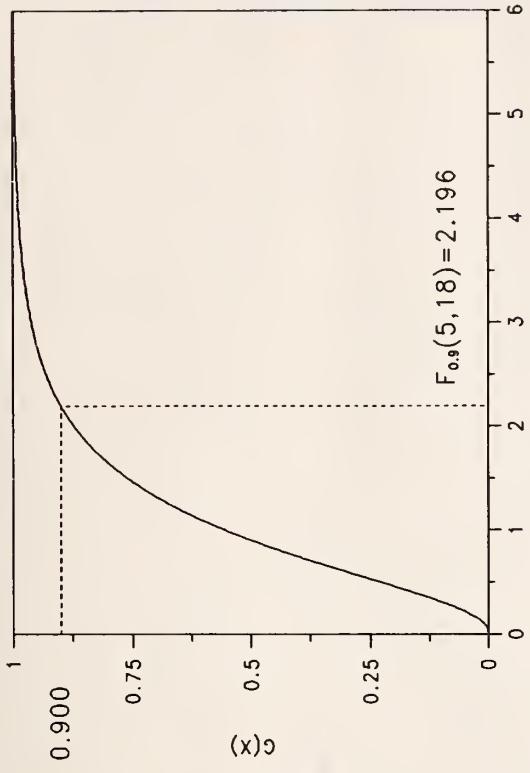
Figure 14

PROBABILITY DENSITY FUNCTION OF t(10) DISTRIBUTION

g(x)

X
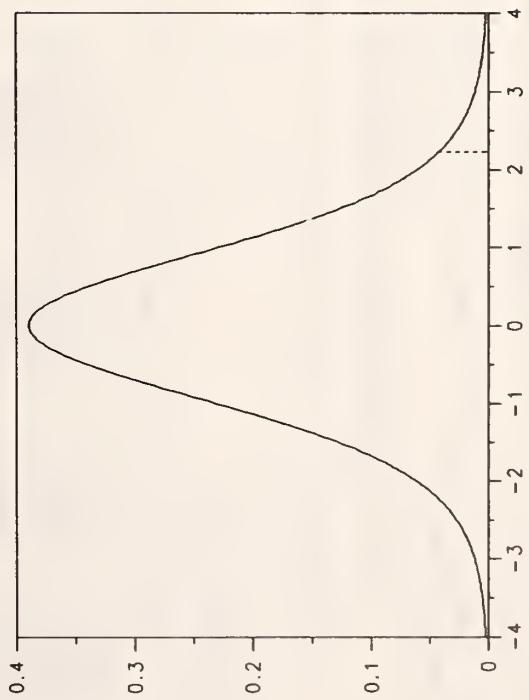
Figure 15

CUMULATIVE DISTRIBUTION FUNCTION OF t(10) DISTRIBUTION

$t_{0.975}(10) = 2.228$

G(x)

0.975

X

Figure 16

31

*********************** PROGRAM FCAL88 FOR FORCE SENSOR CALIBRATION *********************************** PAGE 1
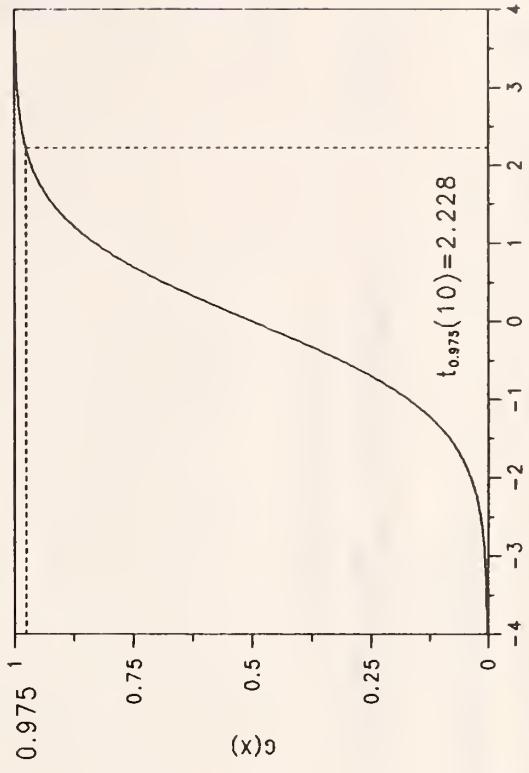
SENSOR: #27662A 737.04/227708 EXAMPLE #8       MACHINE: UNKNOWN (COMPRESSION)
DATE: 05/11/82

REFERENCE: REEVE, CHARLES P., "A NEW STATISTICAL MODEL FOR THE CALIBRATION OF FORCE SENSORS", NATIONAL BUREAU OF STANDARDS, TECHNICAL NOTE 1246, JUNE 1988.

NOTE: LOADS ARE LINEARLY TRANSFORMED TO THE INTERVAL [-1,1] FOR INTERMEDIATE COMPUTATIONS

* SUMMARY OF OBSERVATIONS *

$Y(I,J)$ = OBSERVATION AT LOAD I DURING RUN J

| I | LOAD(I) | Y(I,1) | Y(I,2) | Y(I,3) |
|---|---------|--------|--------|--------|
| 1 | .1000E+05 | .6650000E-01 | .6648000E-01 | .6647000E-01 |
| 2 | .3000E+05 | .1997100 | .1997000 | .1996700 |
| 3 | .6000E+05 | .3997300 | .3997000 | .3996800 |
| 4 | .9000E+05 | .5997500 | .5997900 | .5997800 |
| 5 | .1200E+06 | .7999200 | .7999500 | .7999300 |
| 6 | .1500E+06 | 1.000100 | 1.000120 | 1.000120 |
| 7 | .1800E+06 | 1.200300 | 1.200340 | 1.200330 |
| 8 | .2100E+06 | 1.400500 | 1.400560 | 1.400540 |
| 9 | .2400E+06 | 1.600700 | 1.600780 | 1.600760 |
| 10 | .2700E+06 | 1.800940 | 1.800990 | 1.800980 |
| 11 | .3000E+06 | 2.001100 | 2.001150 | 2.001160 |

* SUMMARY OF MEAN DATA AND DIFFERENCE DATA *

$YBAR(I)$ = MEAN[Y(I,1), Y(I,2), ...]          $Z(I,J)$ = Y(I,J) - YBAR(I)

| I | LOAD(I) | YBAR(I) | Z(I,1) | Z(I,2) | Z(I,3) |
|---|---------|---------|--------|--------|--------|
| 1 | .1000E+05 | .6648333E-01 | .1667E-04 | -.3333E-05 | -.1333E-04 |
| 2 | .3000E+05 | .1996933 | .1667E-04 | .6667E-05 | -.2333E-04 |
| 3 | .6000E+05 | .3997033 | -.2667E-04 | -.3333E-05 | -.2333E-04 |
| 4 | .9000E+05 | .5997733 | -.2333E-04 | .1667E-04 | .6667E-05 |
| 5 | .1200E+06 | .7999333 | -.1333E-04 | .1667E-04 | -.3333E-05 |
| 6 | .1500E+06 | 1.000113 | -.1333E-04 | .6667E-05 | .6667E-05 |
| 7 | .1800E+06 | 1.200323 | -.2333E-04 | .1667E-04 | .6667E-05 |
| 8 | .2100E+06 | 1.400533 | -.3333E-04 | .2667E-04 | .6667E-05 |
| 9 | .2400E+06 | 1.600747 | -.4667E-04 | .3333E-04 | .1333E-04 |
| 10 | .2700E+06 | 1.800970 | -.3000E-04 | .2000E-04 | .1000E-04 |
| 11 | .3000E+06 | 2.001137 | -.3667E-04 | .1333E-04 | .2333E-04 |

PLOT OF DIFFERENCE DATA VS. LOAD (SYMBOL INDICATES RUN NUMBER)
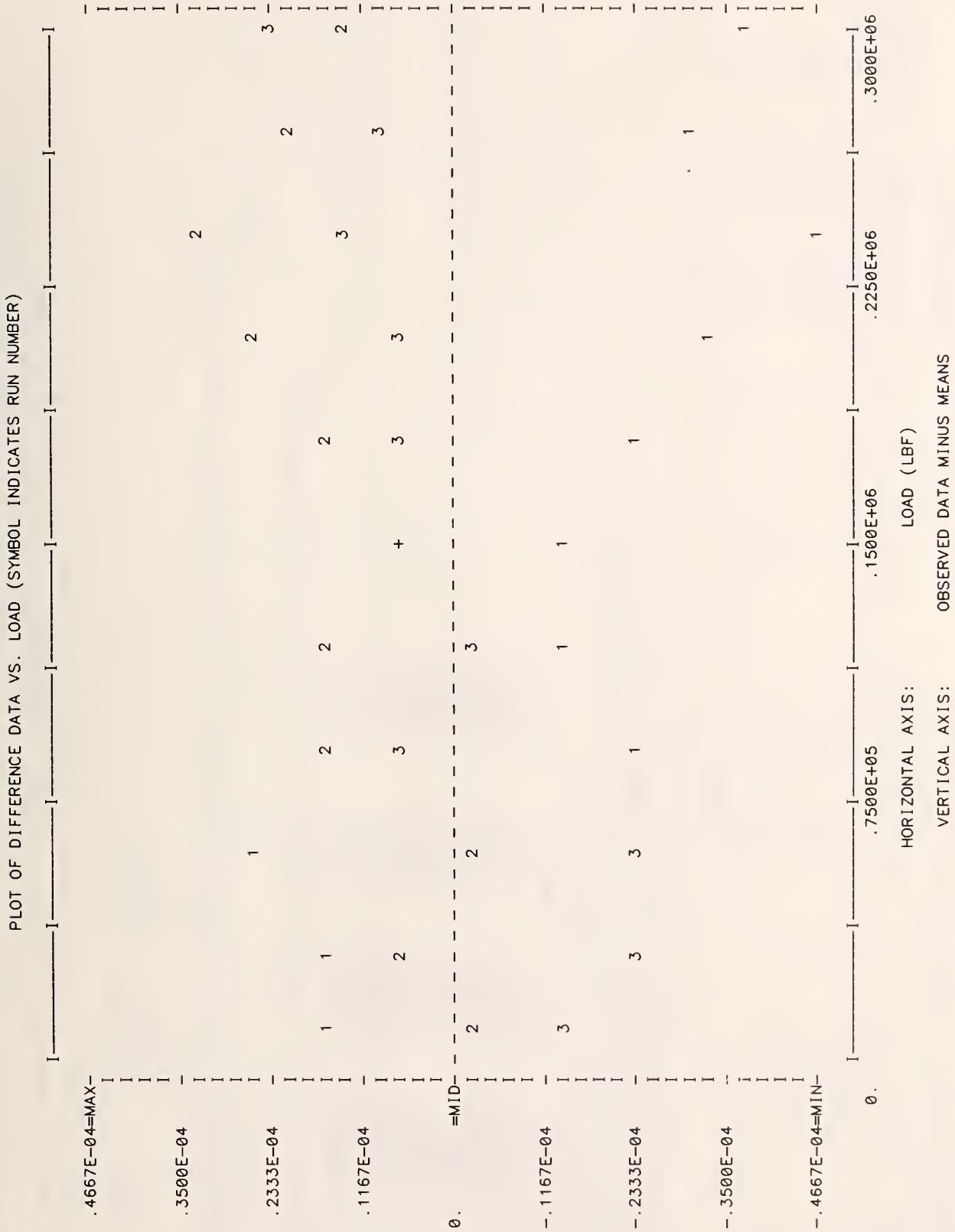
* DETERMINE DEGREE OF POLYNOMIAL FIT TO DIFFERENCE DATA *     MINIMUM DEGREE = -1     MAXIMUM DEGREE = 3

| DEGREE | RESIDUAL STD DEV | DF | TEST OF SIGNIFICANCE OF | F STATISTIC | DF | DF | F C.D.F. | 1-ALPHA | SIGNIFICANT? |
|--------|------------------|-----|-------------------------|-------------|-----|-----|----------|---------|--------------|
| 3 | .119273E-04 | 14. | | | | | | | |
| 2 | .115528E-04 | 16. | CUBIC COEFF | .51 | 2. | 14. | .3862 | .990 | NO |
| 1 | .120239E-04 | 18. | QUADRATIC COEFF | 1.75 | 2. | 16. | .7944 | .990 | NO |
| 0 | .217771E-04 | 20. | LINEAR COEFF | 23.80 | 2. | 18. | 1.0000 | .990 | YES |

### THE BEST FITTING POLYNOMIALS TO THE DIFFERENCE DATA ARE LINEAR

* SUMMARY OF RESIDUALS FROM BEST FITTING POLYNOMIALS TO DIFFERENCE DATA *

$RES(I,J) = Z(I,J) - JTH H*-DEGREE POLYNOMIAL$

| I | LOAD(I) | RES(I,1) | RES(I,2) | RES(I,3) |
|----|-----------|-----------|-----------|-----------|
| 1 | .1000E+05 | .728E-06 | -.549E-05 | .476E-05 |
| 2 | .3000E+05 | .506E-05 | .288E-05 | -.794E-05 |
| 3 | .6000E+05 | .215E-04 | -.956E-05 | -.120E-04 |
| 4 | .9000E+05 | -.220E-04 | .799E-05 | .140E-04 |
| 5 | .1200E+06 | -.547E-05 | .555E-05 | -.744E-07 |
| 6 | .1500E+06 | .102E-05 | -.690E-05 | .588E-05 |
| 7 | .1800E+06 | -.249E-05 | .660E-06 | .183E-05 |
| 8 | .2100E+06 | -.600E-05 | .822E-05 | -.221E-05 |
| 9 | .2400E+06 | -.128E-04 | .124E-04 | .411E-06 |
| 10 | .2700E+06 | .103E-04 | -.334E-05 | -.697E-05 |
| 11 | .3000E+06 | .101E-04 | -.125E-04 | .232E-05 |

RESIDUAL STD DEV = .12024E-04 BASED ON 18. D.F. EXCLUDES EFFECTS OF MODEL ERROR.

* ESTIMATED COVARIANCE MATRIX FOR BETWEEN-RUN EFFECTS *     << BASED ON TRANSFORMED LOADS IN [-1,1] >>

| | CONSTANT | LINEAR |
|----------|-------------|-------------|
| CONSTANT | .217693E-09 | .338829E-09 |
| LINEAR | .338829E-09 | .752960E-09 |

DEGREES OF FREEDOM = 2.

*** BETWEEN-RUN ERROR IS SIGNIFICANT.     FB = 19.48   DF = 4., 18.   C.D.F.(FB) = 1.0000

PLOT OF RESIDUALS FROM H*-DEGREE FIT TO DIFFERENCE DATA VS. LOAD
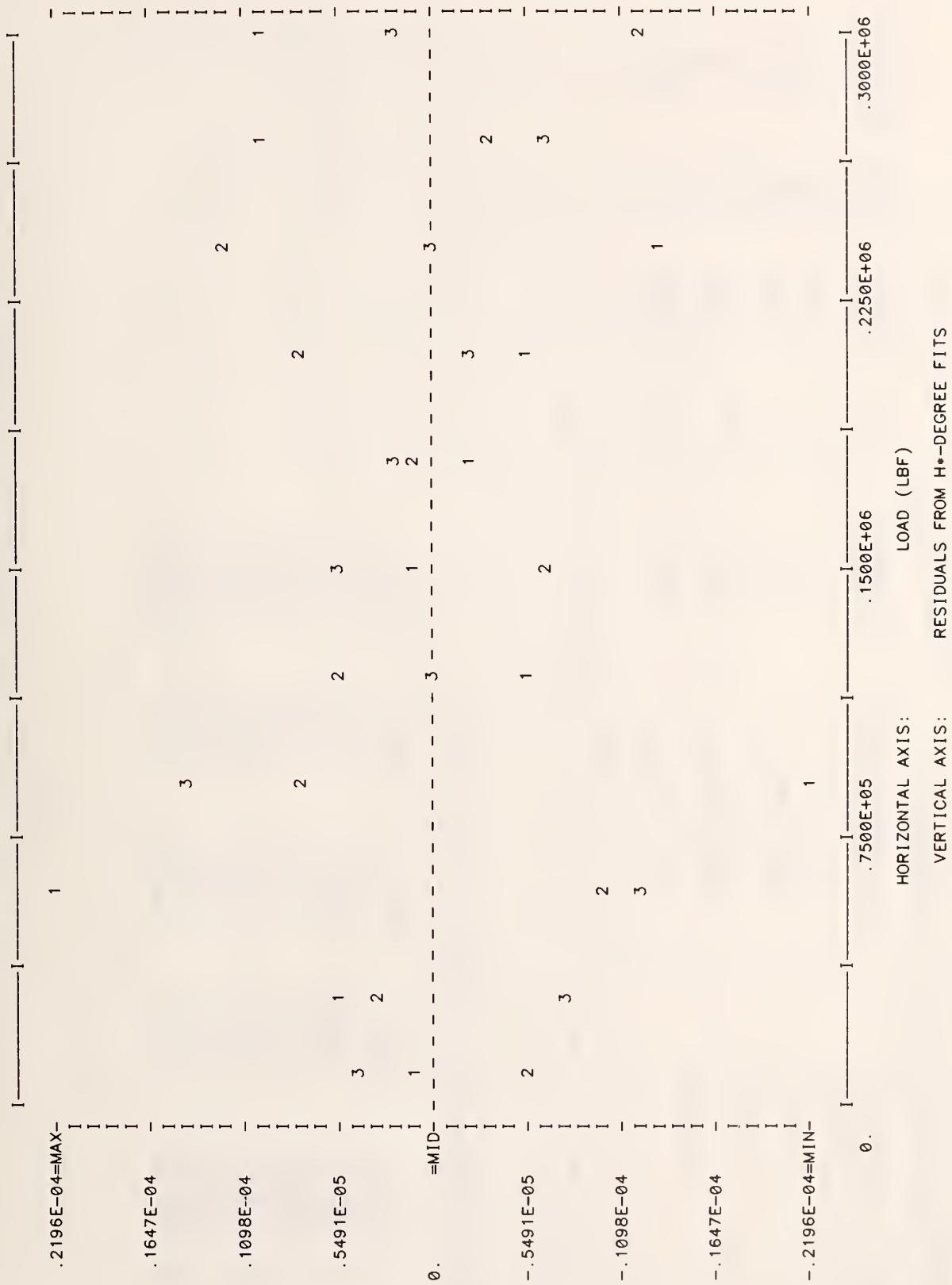
```
.2196E-04=MAX-
              -
              -
.1647E-04     -
              -
.1098E-04     -
              -
.5491E-05     -
              -
0.      =MID- - - - - - - - - - - - - - - - - - - - - - - - -
              -
-.5491E-05    -
              -
-.1098E-04    -
              -
-.1647E-04    -
              -
-.2196E-04=MIN-

              0.        .7500E+05      .1500E+06      .2250E+06      .3000E+06

              HORIZONTAL AXIS:   LOAD (LBF)

              VERTICAL AXIS:     RESIDUALS FROM H*-DEGREE FITS
```

********************** PROGRAM FCAL88 FOR FORCE SENSOR CALIBRATION ************************ PAGE 5

* DETERMINE DEGREE OF POLYNOMIAL FIT TO MEAN DATA *

MINIMUM DEGREE = 1    MAXIMUM DEGREE = 5

| DEGREE | RESIDUAL STD DEV | DF | TEST OF SIGNIFICANCE OF | F STATISTIC | DF | DF | F C.D.F. | 1-ALPHA | SIGNIFICANT? |
|---|---|---|---|---|---|---|---|---|---|
| 5 | .102040E-04 | 5. | MODEL ERROR | 2.16 | 5. | 18. | .8956 | .900 | NO |
| 4 | .146005E-04 | 6. | MODEL ERROR | 4.42 | 6. | 18. | .9936 | .900 | YES |
|  |  |  | QUINTIC COEFF | 7.28 | 1. | 5. | .9572 | .975 | NO |
| 3 | .204370E-04 | 7. | MODEL ERROR | 8.67 | 7. | 18. | .9999 | .900 | YES |
|  |  |  | QUARTIC COEFF | 7.71 | 1. | 6. | .9679 | .975 | NO |
| 2 | .754384E-04 | 8. | MODEL ERROR | 118.09 | 8. | 18. | 1.0000 | .900 | YES |
|  |  |  | CUBIC COEFF | 102.00 | 1. | 7. | 1.0000 | .975 | YES |

### THE BEST FITTING POLYNOMIAL TO MEAN DATA IS A CUBIC

RESIDUAL STD DEV = .20437E-04  BASED ON  7.  D.F.  INCLUDES EFFECTS OF MODEL ERROR.

*** MODEL ERROR IS SIGNIFICANT.

SUMMARY OF RESIDUALS FROM POLYNOMIAL FITS TO MEAN DATA *

| I | LOAD(I) | QUADRATIC RES(I) | CUBIC RES(I) | QUARTIC RES(I) | QUINTIC RES(I) |
|---|---|---|---|---|---|
| 1 | .1000E+05 | .115E-03 | .269E-04 | .119E-04 | .353E-05 |
| 2 | .3000E+05 | -.288E-04 | -.320E-04 | -.224E-04 | -.108E-04 |
| 3 | .6000E+05 | -.729E-04 | -.785E-05 | .881E-05 | -.147E-04 |
| 4 | .9000E+05 | -.853E-04 | -.731E-05 | -.133E-05 | -.836E-05 |
| 5 | .1200E+06 | -.359E-04 | .173E-04 | .972E-05 | .232E-07 |
| 6 | .1500E+06 | .539E-05 | .134E-04 | -.106E-05 | -.278E-05 |
| 7 | .1800E+06 | .484E-04 | .876E-05 | -.230E-05 | .575E-05 |
| 8 | .2100E+06 | .633E-04 | -.913E-05 | -.824E-05 | .119E-05 |
| 9 | .2400E+06 | .533E-04 | -.193E-04 | -.537E-05 | -.665E-05 |
| 10 | .2700E+06 | .252E-04 | .237E-05 | .175E-04 | -.422E-05 |
| 11 | .3000E+06 | -.879E-04 | .692E-05 | -.721E-05 | -.858E-06 |

PLOT OF RESIDUALS FROM [SYMBOL]-DEGREE POLYNOMIAL FIT TO MEAN DATA



HORIZONTAL AXIS:     LOAD (LBF)

VERTICAL AXIS:     RESIDUALS FROM POLYNOMIAL FITS

37
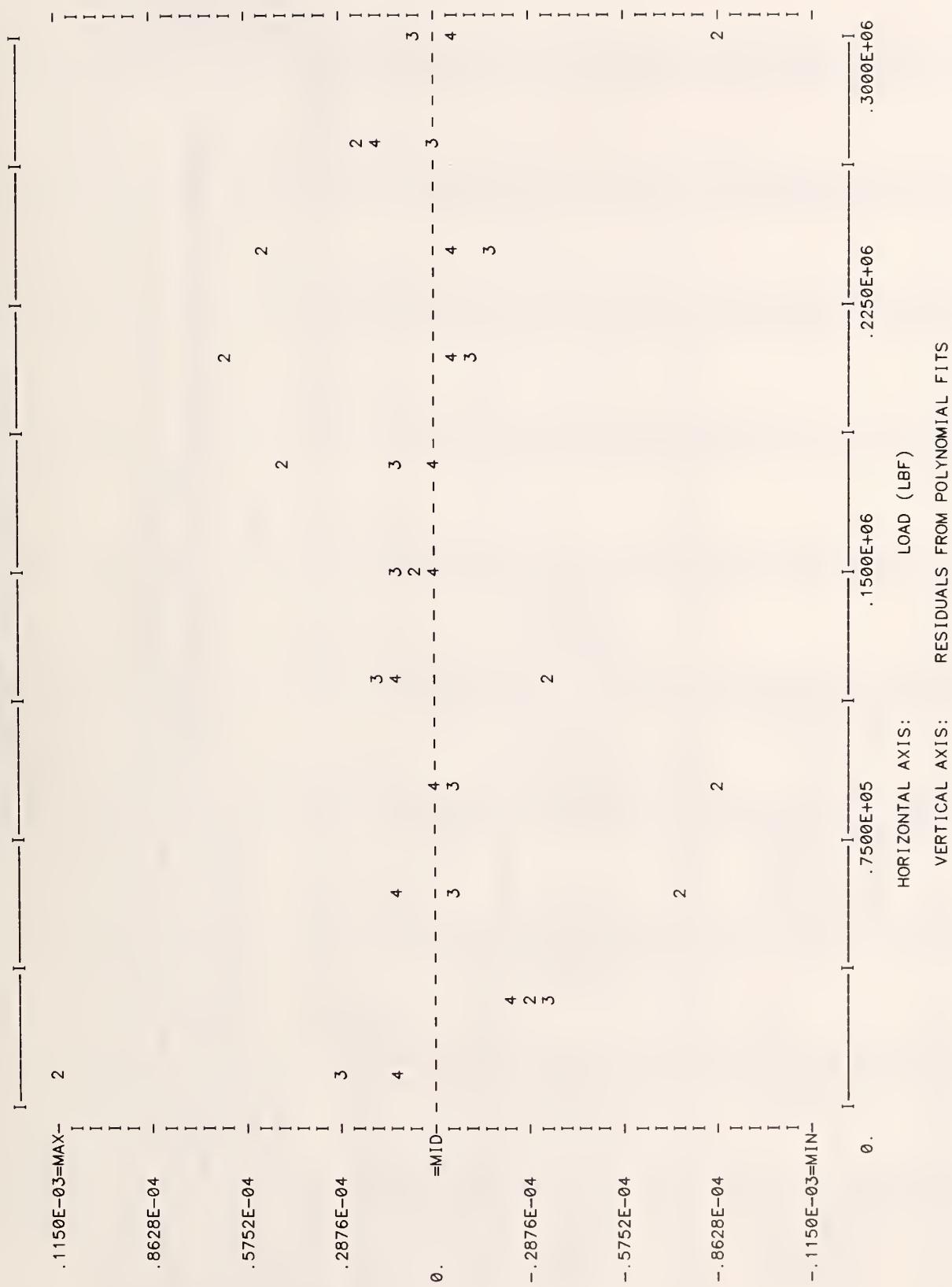
* ESTIMATES OF POLYNOMIAL RESPONSE CURVE COEFFICIENTS FROM MEAN DATA *     << BASED ON TRANSFORMED LOADS IN [-1,1] >>

| COEFFICIENT | ESTIMATE | STANDARD DEVIATION | T-VALUE EST/STD DEV |
|---|---|---|---|
| THETA0 | 1.033467 | .9451697E-05 | 109341.98 |
| THETA1 | .9676675 | .2556314E-04 | 37854.02 |
| THETA2 | .3257531E-03 | .1724280E-04 | 18.89 |
| THETA3 | -.3308707E-03 | .3276050E-04 | -10.10 |

* ESTIMATED COVARIANCE MATRIX FOR POLYNOMIAL COEFFICIENTS *      << BASED ON TRANSFORMED LOADS IN [-1,1] >>

| | CONSTANT | LINEAR | QUADRATIC | CUBIC |
|---|---|---|---|---|
| CONSTANT | .893346E-10 | -.623363E-12 | -.123472E-09 | -.685309E-12 |
| LINEAR | -.623363E-12 | .653474E-09 | .667837E-12 | -.776514E-09 |
| QUADRATIC | -.123472E-09 | .667837E-12 | .297314E-09 | .113785E-10 |
| CUBIC | -.685309E-12 | -.776514E-09 | .113785E-10 | .107325E-08 |

THE VARIANCE ESTIMATE WHICH SCALES THE ABOVE MATRIX IS DETERMINED FROM THE MEAN DATA ONLY
SINCE MODEL ERROR IS SIGNIFICANT.

VARIANCE = .41767E-09     STD DEV = .20437E-04     D.F. = 7.

* SUMMARY OF ERROR CONTRIBUTIONS *

| SOURCE OF ERROR | RELATIVE SIZES OF "SIGMA-LIKE" VALUES OF CONTRIBUTING ERROR IN THIS CALIBRATION | "SIGMA" |
|---|---|---|
| : ROUNDOFF | :RRRRRRRRR | .2887E-05 |
| : WITHIN-RUN | :EEEEEEEEEEEEEEEEEEEEEEEEEE | .1167E-04 |
| : MODEL | :MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM | .3329E-04 |
| : BETWEEN-RUN | :BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB | .2263E-04 |
| : SYSTEMATIC | :SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS | .2013E-04 |

CONFIDENCE AND PREDICTION INTERVALS / INVERSE PREDICTION *

IN COLUMNS 1,2,3,4,5,6,7,8    DELTA = .050
IN COLUMNS 7,8                ALPHA = .050

COLUMNS 1,2,3,4,5,6 MAY BE USED FOR CONFIDENCE AND/OR PREDICTION INTERVAL APPLICATIONS

COLUMNS 3,4 MAY BE USED FOR "SINGLE USE" INVERSE PREDICTION ON THE SAME MACHINE
COLUMNS 5,6 MAY BE USED FOR "MULTIPLE USE" INVERSE PREDICTION ON A DIFFERENT MACHINE
COLUMNS 7,8 MAY BE USED FOR "MULTIPLE USE" INVERSE PREDICTION ON THE SAME MACHINE

(THESE USES ARE DESCRIBED IN THE REFERENCE ON PAGE 1 OF THIS PRINTOUT.)

| LOAD | PREDICTED RESPONSE | [1] CONFIDENCE INTERVAL FOR SINGLE FUNC VALUE (LOWER) | [2] CONFIDENCE INTERVAL FOR SINGLE FUNC VALUE (UPPER) | [3] PREDICTION INTERVAL FOR NEW OBSERVATION (LOWER) | [4] PREDICTION INTERVAL FOR NEW OBSERVATION (UPPER) | [5] CONFIDENCE INTERVAL FOR ENTIRE FUNCTION (LOWER) | [6] CONFIDENCE INTERVAL FOR ENTIRE FUNCTION (UPPER) | [7] COMBINED INTERVALS FOR ENTIRE FUNCTION AND NEW OBSERVATION (LOWER) | [8] COMBINED INTERVALS FOR ENTIRE FUNCTION AND NEW OBSERVATION (UPPER) |
|---|---|---|---|---|---|---|---|---|---|
| 12000.0 | .797812E-01 | .797383E-01 | .798241E-01 | .797159E-01 | .798465E-01 | .797089E-01 | .798535E-01 | .796535E-01 | .799089E-01 |
| 15000.0 | .997692E-01 | .997288E-01 | .998096E-01 | .997062E-01 | .998322E-01 | .997010E-01 | .998373E-01 | .996469E-01 | .998914E-01 |
| 18000.0 | .119758 | .119720 | .119796 | .119697 | .119819 | .119694 | .119823 | .119641 | .119876 |
| 21000.0 | .139748 | .139712 | .139785 | .139690 | .139807 | .139687 | .139810 | .139636 | .139861 |
| 24000.0 | .159740 | .159705 | .159774 | .159683 | .159797 | .159681 | .159798 | .159631 | .159849 |
| 27000.0 | .179732 | .179699 | .179765 | .179677 | .179787 | .179676 | .179788 | .179627 | .179837 |
| 30000.0 | .199725 | .199694 | .199757 | .199672 | .199779 | .199671 | .199779 | .199623 | .199827 |
| 33000.0 | .219720 | .219689 | .219750 | .219668 | .219772 | .219668 | .219772 | .219620 | .219819 |
| 36000.0 | .239715 | .239685 | .239745 | .239665 | .239765 | .239664 | .239765 | .239618 | .239812 |
| 39000.0 | .259711 | .259682 | .259740 | .259662 | .259760 | .259662 | .259760 | .259616 | .259806 |
| 42000.0 | .279708 | .279680 | .279737 | .279661 | .279756 | .279660 | .279757 | .279616 | .279801 |
| 45000.0 | .299707 | .299679 | .299735 | .299660 | .299753 | .299659 | .299754 | .299615 | .299798 |
| 48000.0 | .319706 | .319678 | .319734 | .319660 | .319751 | .319659 | .319753 | .319616 | .319796 |
| 51000.0 | .339706 | .339678 | .339733 | .339661 | .339751 | .339659 | .339752 | .339617 | .339795 |
| 54000.0 | .359707 | .359679 | .359734 | .359663 | .359751 | .359660 | .359753 | .359618 | .359795 |
| 57000.0 | .379709 | .379681 | .379736 | .379665 | .379752 | .379662 | .379755 | .379621 | .379796 |
| 60000.0 | .399711 | .399684 | .399739 | .399669 | .399754 | .399665 | .399758 | .399624 | .399798 |
| 63000.0 | .419715 | .419687 | .419742 | .419673 | .419757 | .419668 | .419761 | .419628 | .419801 |
| 66000.0 | .439719 | .439691 | .439747 | .439677 | .439760 | .439672 | .439766 | .439632 | .439806 |
| 69000.0 | .459724 | .459696 | .459752 | .459683 | .459765 | .459677 | .459771 | .459638 | .459810 |
| 72000.0 | .479730 | .479702 | .479758 | .479689 | .479771 | .479683 | .479777 | .479644 | .479816 |
| 75000.0 | .499737 | .499709 | .499764 | .499696 | .499777 | .499690 | .499784 | .499650 | .499823 |
| 78000.0 | .519744 | .519716 | .519772 | .519704 | .519784 | .519697 | .519791 | .519658 | .519830 |
| 81000.0 | .539752 | .539724 | .539780 | .539712 | .539792 | .539705 | .539799 | .539666 | .539838 |
| 84000.0 | .559761 | .559733 | .559789 | .559721 | .559801 | .559714 | .559808 | .559675 | .559847 |
| 87000.0 | .579770 | .579742 | .579798 | .579730 | .579810 | .579723 | .579818 | .579685 | .579856 |
| 90000.0 | .599781 | .599753 | .599809 | .599741 | .599821 | .599733 | .599828 | .599695 | .599866 |
| 93000.0 | .619792 | .619764 | .619819 | .619752 | .619831 | .619744 | .619839 | .619706 | .619877 |
| 96000.0 | .639803 | .639775 | .639831 | .639763 | .639843 | .639756 | .639850 | .639717 | .639889 |
| 99000.0 | .659815 | .659787 | .659843 | .659775 | .659855 | .659768 | .659862 | .659730 | .659901 |
| 102000. | .679828 | .679800 | .679855 | .679788 | .679868 | .679781 | .679874 | .679742 | .679913 |
| 105000. | .699841 | .699814 | .699869 | .699801 | .699882 | .699795 | .699888 | .699756 | .699927 |
| 108000. | .719855 | .719828 | .719882 | .719814 | .719896 | .719809 | .719901 | .719770 | .719941 |
| 111000. | .739870 | .739842 | .739897 | .739829 | .739911 | .739824 | .739915 | .739784 | .739955 |
| 114000. | .759885 | .759857 | .759912 | .759843 | .759926 | .759839 | .759930 | .759799 | .759970 |
| 117000. | .779900 | .779873 | .779927 | .779858 | .779942 | .779855 | .779946 | .779814 | .779986 |
| 120000. | .799916 | .799889 | .799943 | .799874 | .799958 | .799871 | .799961 | .799830 | .800002 |

B-8

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 123000. | .819933 | .819906 | .819959 | .819890 | .819976 | .819887 | .819978 | .819846 | .820019 |
| 126000. | .839950 | .839923 | .839976 | .839906 | .839993 | .839904 | .839995 | .839862 | .840037 |
| 129000. | .859967 | .859940 | .859994 | .859922 | .860011 | .859921 | .860013 | .859879 | .860055 |
| 132000. | .879985 | .879958 | .880012 | .879939 | .880030 | .879939 | .880031 | .879896 | .880074 |
| 135000. | .900003 | .899976 | .900030 | .899957 | .900050 | .899957 | .900049 | .899913 | .900093 |
| 138000. | .920022 | .919994 | .920049 | .919974 | .920069 | .919975 | .920068 | .919930 | .920113 |
| 141000. | .940041 | .940013 | .940069 | .939992 | .940090 | .939993 | .940088 | .939948 | .940134 |
| 144000. | .960060 | .960032 | .960088 | .960010 | .960110 | .960012 | .960108 | .959966 | .960155 |
| 147000. | .980080 | .980051 | .980109 | .980028 | .980131 | .980031 | .980129 | .979984 | .980176 |
| 150000. | 1.00010 | 1.00007 | 1.00013 | 1.00005 | 1.00015 | 1.00005 | 1.00015 | 1.00000 | 1.00020 |
| 153000. | 1.02012 | 1.02009 | 1.02015 | 1.02007 | 1.02017 | 1.02007 | 1.02017 | 1.02002 | 1.02022 |
| 156000. | 1.04014 | 1.04011 | 1.04017 | 1.04008 | 1.04019 | 1.04009 | 1.04019 | 1.04004 | 1.04024 |
| 159000. | 1.06016 | 1.06013 | 1.06019 | 1.06010 | 1.06022 | 1.06011 | 1.06022 | 1.06006 | 1.06027 |
| 162000. | 1.08018 | 1.08015 | 1.08021 | 1.08012 | 1.08024 | 1.08013 | 1.08024 | 1.08008 | 1.08029 |
| 165000. | 1.10020 | 1.10017 | 1.10024 | 1.10014 | 1.10027 | 1.10015 | 1.10026 | 1.10009 | 1.10032 |
| 168000. | 1.12023 | 1.12019 | 1.12026 | 1.12016 | 1.12029 | 1.12017 | 1.12028 | 1.12011 | 1.12034 |
| 171000. | 1.14025 | 1.14021 | 1.14028 | 1.14018 | 1.14031 | 1.14019 | 1.14031 | 1.14013 | 1.14036 |
| 174000. | 1.16027 | 1.16024 | 1.16031 | 1.16020 | 1.16034 | 1.16021 | 1.16033 | 1.16015 | 1.16039 |
| 177000. | 1.18029 | 1.18026 | 1.18033 | 1.18022 | 1.18036 | 1.18023 | 1.18035 | 1.18017 | 1.18041 |
| 180000. | 1.20031 | 1.20028 | 1.20035 | 1.20024 | 1.20039 | 1.20025 | 1.20038 | 1.20019 | 1.20044 |
| 183000. | 1.22034 | 1.22030 | 1.22037 | 1.22026 | 1.22041 | 1.22027 | 1.22040 | 1.22021 | 1.22047 |
| 186000. | 1.24036 | 1.24032 | 1.24040 | 1.24028 | 1.24044 | 1.24029 | 1.24043 | 1.24023 | 1.24049 |
| 189000. | 1.26038 | 1.26034 | 1.26042 | 1.26030 | 1.26046 | 1.26031 | 1.26045 | 1.26025 | 1.26052 |
| 192000. | 1.28041 | 1.28036 | 1.28045 | 1.28032 | 1.28049 | 1.28033 | 1.28048 | 1.28027 | 1.28054 |
| 195000. | 1.30043 | 1.30039 | 1.30047 | 1.30035 | 1.30051 | 1.30035 | 1.30050 | 1.30028 | 1.30057 |
| 198000. | 1.32045 | 1.32041 | 1.32049 | 1.32037 | 1.32054 | 1.32038 | 1.32053 | 1.32030 | 1.32060 |
| 201000. | 1.34047 | 1.34043 | 1.34052 | 1.34039 | 1.34056 | 1.34040 | 1.34055 | 1.34032 | 1.34062 |
| 204000. | 1.36050 | 1.36045 | 1.36054 | 1.36041 | 1.36059 | 1.36042 | 1.36058 | 1.36034 | 1.36065 |
| 207000. | 1.38052 | 1.38047 | 1.38057 | 1.38043 | 1.38061 | 1.38044 | 1.38060 | 1.38036 | 1.38068 |
| 210000. | 1.40054 | 1.40050 | 1.40059 | 1.40045 | 1.40064 | 1.40046 | 1.40063 | 1.40038 | 1.40070 |
| 213000. | 1.42057 | 1.42052 | 1.42061 | 1.42047 | 1.42066 | 1.42048 | 1.42065 | 1.42040 | 1.42073 |
| 216000. | 1.44059 | 1.44054 | 1.44064 | 1.44049 | 1.44069 | 1.44050 | 1.44067 | 1.44042 | 1.44076 |
| 219000. | 1.46061 | 1.46056 | 1.46066 | 1.46051 | 1.46071 | 1.46052 | 1.46070 | 1.46044 | 1.46078 |
| 222000. | 1.48063 | 1.48058 | 1.48068 | 1.48053 | 1.48074 | 1.48054 | 1.48072 | 1.48046 | 1.48081 |
| 225000. | 1.50066 | 1.50060 | 1.50071 | 1.50055 | 1.50076 | 1.50056 | 1.50075 | 1.50047 | 1.50084 |
| 228000. | 1.52068 | 1.52063 | 1.52073 | 1.52057 | 1.52079 | 1.52058 | 1.52077 | 1.52049 | 1.52086 |
| 231000. | 1.54070 | 1.54065 | 1.54075 | 1.54059 | 1.54081 | 1.54060 | 1.54080 | 1.54051 | 1.54089 |
| 234000. | 1.56072 | 1.56067 | 1.56078 | 1.56061 | 1.56084 | 1.56062 | 1.56082 | 1.56053 | 1.56092 |
| 237000. | 1.58074 | 1.58069 | 1.58080 | 1.58063 | 1.58086 | 1.58064 | 1.58085 | 1.58054 | 1.58094 |
| 240000. | 1.60077 | 1.60071 | 1.60082 | 1.60065 | 1.60089 | 1.60066 | 1.60087 | 1.60056 | 1.60097 |
| 243000. | 1.62079 | 1.62073 | 1.62085 | 1.62066 | 1.62091 | 1.62068 | 1.62089 | 1.62058 | 1.62100 |
| 246000. | 1.64081 | 1.64075 | 1.64087 | 1.64068 | 1.64093 | 1.64070 | 1.64092 | 1.64059 | 1.64102 |
| 249000. | 1.66083 | 1.66077 | 1.66089 | 1.66070 | 1.66096 | 1.66072 | 1.66094 | 1.66061 | 1.66105 |
| 252000. | 1.68085 | 1.68079 | 1.68091 | 1.68072 | 1.68098 | 1.68073 | 1.68097 | 1.68063 | 1.68107 |
| 255000. | 1.70087 | 1.70081 | 1.70093 | 1.70074 | 1.70101 | 1.70075 | 1.70099 | 1.70064 | 1.70110 |
| 258000. | 1.72089 | 1.72083 | 1.72096 | 1.72075 | 1.72103 | 1.72077 | 1.72101 | 1.72066 | 1.72113 |
| 261000. | 1.74091 | 1.74084 | 1.74098 | 1.74077 | 1.74105 | 1.74079 | 1.74103 | 1.74067 | 1.74115 |
| 264000. | 1.76093 | 1.76086 | 1.76100 | 1.76079 | 1.76107 | 1.76080 | 1.76106 | 1.76069 | 1.76117 |
| 267000. | 1.78095 | 1.78088 | 1.78102 | 1.78080 | 1.78109 | 1.78082 | 1.78108 | 1.78070 | 1.78120 |
| 270000. | 1.80097 | 1.80090 | 1.80104 | 1.80082 | 1.80112 | 1.80084 | 1.80110 | 1.80071 | 1.80122 |
| 273000. | 1.82099 | 1.82091 | 1.82106 | 1.82084 | 1.82114 | 1.82085 | 1.82112 | 1.82073 | 1.82124 |
| 276000. | 1.84100 | 1.84093 | 1.84108 | 1.84085 | 1.84116 | 1.84087 | 1.84114 | 1.84074 | 1.84127 |
| 279000. | 1.86102 | 1.86095 | 1.86109 | 1.86087 | 1.86118 | 1.86089 | 1.86116 | 1.86076 | 1.86129 |
| 282000. | 1.88104 | 1.88096 | 1.88111 | 1.88088 | 1.88120 | 1.88091 | 1.88117 | 1.88077 | 1.88131 |
| 285000. | 1.90105 | 1.90098 | 1.90113 | 1.90090 | 1.90121 | 1.90092 | 1.90119 | 1.90078 | 1.90133 |
| 288000. | 1.92107 | 1.92100 | 1.92115 | 1.92091 | 1.92123 | 1.92093 | 1.92121 | 1.92080 | 1.92135 |
| 291000. | 1.94109 | 1.94101 | 1.94116 | 1.94092 | 1.94125 | 1.94095 | 1.94122 | 1.94081 | 1.94136 |
| 294000. | 1.96110 | 1.96103 | 1.96118 | 1.96094 | 1.96127 | 1.96096 | 1.96124 | 1.96082 | 1.96138 |
| 297000. | 1.98112 | 1.98104 | 1.98119 | 1.98095 | 1.98128 | 1.98098 | 1.98125 | 1.98083 | 1.98140 |
| 300000. | 2.00113 | 2.00105 | 2.00121 | 2.00096 | 2.00130 | 2.00099 | 2.00127 | 2.00084 | 2.00142 |

95% C.I. FOR SINGLE POINT ON CURVE (S), ENTIRE CURVE (E), & NEW OBS (N)

```
.1664E-03=MAX
.1248E-03
.8318E-04
.4159E-04
0.         =MID
-.4159E-04
-.8318E-04
-.1248E-03
-.1664E-03=MIN

            0.        .7500E+05     .1500E+06     .2250E+06     .3000E+06

            HORIZONTAL AXIS:   LOAD (LBF)

            VERTICAL AXIS:     CONF LIMITS ABOUT PREDICTED CURVE
```

| U.S. DEPT. OF COMM.<br><br>**BIBLIOGRAPHIC DATA**<br>**SHEET** (See instructions) | 1. PUBLICATION OR<br>REPORT NO.<br><br>NBS/TN-1246 | 2. Performing Organ. Report No. | 3. Publication Date<br><br>June 1988 |
|---|---|---|---|

**4. TITLE AND SUBTITLE**

A New Statistical Model for the Calibration of Force Sensors

**5. AUTHOR(S)**

Charles P. Reeve

| **6. PERFORMING ORGANIZATION** (If joint or other than NBS, see instructions)<br><br>**NATIONAL BUREAU OF STANDARDS**<br>**U.S. DEPARTMENT OF COMMERCE**<br>**GAITHERSBURG, MD 20899** | 7. Contract/Grant No. |
|---|---|
| | 8. Type of Report & Period Covered<br><br>Final |

**9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS** (Street, City, State, ZIP)

Same as #6

**10. SUPPLEMENTARY NOTES**

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

**11. ABSTRACT** (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

The National Bureau of Standards has been calibrating force sensors for many years. The objective in these calibrations is to determine the functional relationship between the applied load and sensor response. In a typical calibration several runs are made in which identical sequences of known loads are applied to the sensor. The sensor is rotated in the loading machine between runs. The previous method of analysis incorporated a quadratic polynomial model which was fit to the pooled data. The new method presented here fits separate polynomials to the mean data and between-run differences. The "best" degrees of these polynomials are automatically determined by algorithms which incorporate statistical tests. As a result, error contributions from several sources are quantized. Methods for computing confidence intervals for the "true" sensor response and for a new observation are given, and methods of inverse prediction (measurement of an unknown load) based on these intervals are illustrated.

**12. KEY WORDS** (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

calibration curve; confidence interval; force sensor; inverse prediction; least squares polynomial; load cell; proving ring; statistical model

**13. AVAILABILITY**

☒ Unlimited

☐ For Official Distribution. Do Not Release to NTIS

☒ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

Order From National Technical Information Service (NTIS), Springfield, VA. 22161

**14. NO. OF PRINTED PAGES**

46

**15. Price**

# NBS Technical Publications

## Periodical

**Journal of Research**—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. Issued six times a year.

## Nonperiodicals

**Monographs**—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

**Handbooks**—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

**Special Publications**—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

**Applied Mathematics Series**—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

**National Standard Reference Data Series**—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements are available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

**Building Science Series**—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

**Technical Notes**—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

**Voluntary Product Standards**—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

**Consumer Information Series**—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

*Order the* **above** *NBS publications from: Superintendent of Documents, Government Printing Office. Washington, DC 20402.*

*Order the* **following** *NBS publications—FIPS and NBSIR's—from the National Technical Information Service, Springfield, VA 22161.*

**Federal Information Processing Standards Publications (FIPS PUB)**—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

**NBS Interagency Reports (NBSIR)**—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161, in paper copy or microfiche form.

75 Years Stimulating America's Progress
1913-1988